

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:43:12 : Search time 3799 Seconds
(without alignments)
9015.042 Million cell updates/sec

Title: US-10-715-129-1

Perfect score: 732

Sequence: 1 tctctcccccattcccca.....gaattatcatatgatgat 732

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. And is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|--------------------|
| 1 | 531.4 | 72.6 | 543 | 1 | A1998680 |
| 2 | 430 | 58.7 | 478 | 11 | CR356994 Arabidops |
| 3 | 401 | 54.8 | 477 | 9 | BH753057 |
| 4 | 364.6 | 49.8 | 385 | 3 | BP851800 |
| 5 | 262.4 | 35.8 | 920 | 10 | CL485455 |
| 6 | 239 | 32.7 | 639 | 9 | CC963171 |
| 7 | 191 | 26.1 | 321 | 9 | BH904895 |
| 8 | 191 | 26.1 | 371 | 9 | BH904894 |
| 9 | 182.6 | 24.9 | 471 | 1 | AV544303 |
| 10 | 181 | 24.7 | 763 | 4 | CNS0A8DR |
| 11 | 140.8 | 19.2 | 616 | 9 | BH528419 |
| 12 | 118 | 16.1 | 850 | 4 | CNS0A7C |
| 13 | 114.8 | 15.7 | 793 | 9 | BZ039389 |
| 14 | 102.8 | 14.0 | 474 | 7 | CV432287 |
| 15 | 99.6 | 13.6 | 765 | 9 | CC737187 |
| 16 | 98 | 13.4 | 585 | 9 | CC737307 |
| 17 | 96.4 | 13.2 | 669 | 9 | CC737238 |
| 18 | 96.4 | 13.2 | 904 | 6 | CA994079 |
| 19 | 93.2 | 12.7 | 646 | 6 | CC737303 |
| 20 | 91.2 | 12.5 | 679 | 3 | BH578473 |
| 21 | 88 | 12.0 | 621 | 9 | CC737288 |
| 22 | 86.6 | 11.8 | 369 | 7 | CNS48499 |

| | | | | | |
|------|------|------|-----|----|----------|
| c 23 | 85.4 | 11.7 | 291 | 10 | BX657886 |
| c 24 | 85.2 | 11.6 | 622 | 7 | CN893349 |
| c 25 | 85 | 11.6 | 578 | 6 | CR404276 |
| c 26 | 83.2 | 11.4 | 595 | 9 | CC737194 |
| c 27 | 83.2 | 11.4 | 602 | 9 | CC737269 |
| c 28 | 82.8 | 11.3 | 382 | 2 | BG125370 |
| c 29 | 82.8 | 11.3 | 624 | 1 | A1774560 |
| c 30 | 82.6 | 11.3 | 606 | 3 | B1934590 |
| c 31 | 82.4 | 11.3 | 694 | 7 | CV255174 |
| c 32 | 82.4 | 11.3 | 771 | 7 | CV266016 |
| c 33 | 82.4 | 11.3 | 817 | 7 | CV243132 |
| c 34 | 80.8 | 11.0 | 758 | 8 | CK655014 |
| c 35 | 79.8 | 10.9 | 598 | 8 | CK175107 |
| c 36 | 79.8 | 10.9 | 635 | 7 | CN193375 |
| c 37 | 79.8 | 10.9 | 657 | 7 | CN192376 |
| c 38 | 79.8 | 10.9 | 661 | 7 | CV476718 |
| c 39 | 79.8 | 10.9 | 678 | 8 | DN587342 |
| c 40 | 79.8 | 10.9 | 749 | 8 | DN625498 |
| c 41 | 79.8 | 10.9 | 752 | 7 | CF831325 |
| c 42 | 79.8 | 10.9 | 775 | 7 | CN187913 |
| c 43 | 79.8 | 10.9 | 779 | 7 | CN188780 |
| c 44 | 79.8 | 10.9 | 859 | 8 | CK045872 |
| c 45 | 79.6 | 10.9 | 678 | 5 | BW688619 |

ALIGNMENTS

RESULT 1
LOCUS A1998680/c 543 bp mRNA linear EST 08-SEP-1999
DEFINITION 701546552 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
thaliana cDNA clone 701546552, mRNA sequence.
ACCESSION A1998680
VERSION A1998680.1 GI:5845585
KEYWORDS EST.

ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana

REFERENCE
AUTHORS Chen, J., Montlyama, M., Chan, B., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzozka, P., Gorgone, G., Burns, D., Griffin, J., Mounoutova, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kaestury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.

TITLE Arabidopsis thaliana Gene Expression Microarray
JOURNAL Unpublished (1999)
COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.
Location/Qualifiers

FEATURES
SOURCE 1. 543
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="701546552"
/issue_type="rosette"
/dev_str="4 - 7 weeks"
/clone_id="A. thaliana, Columbia Col-0, rosette-2"
/note="Vector: pSPORT, Site_1: NotI, Site_2: SalI, cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and

watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adapters, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN

Query Match 72.6%; Score 531.4; DB 1; Length 543;
Best Local Similarity 98.5%; Pred. No. 1,7e-135;
Matches 532; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 192 CTTTACCAATGCTTCATCAAGCTGATGTTACCAAGTGGCTTAAGATCAGATCT 251
Db CTTTACCAATGCTTCATCAAGCTGATGTTACCAAGTGGCTTAAGATCAGATCT 484
QY 252 TTGAGCAAAATTTTGTACATGATCTTTCTAAATCCAAACGACACACAAACCTCT 311
Db TTGAGCAAAATTTTGTACATGATCTTTCTAAATCCAAACGACACACAAACCTCT 424
QY 312 GTAGATGATTCGAGAAATCTCAACTTACAAAAGATATTAAACATTCAGACAGTTA 371
Db GTAGATGATTCGAGAAATCTCAACTTACAAAAGATATTAAACATTCAGACAGTTA 364
QY 372 TTGCAACACCGAGTCATGACGTGCGAAGAAACACCGAAGAAACATGAGCTTCGAG 431
Db TTGCAACACCGAGTCATGACGTGCGAAGAAACACCGAAGAAACATGAGCTTCGAG 304
QY 432 TTGCGCGGAGAAAAGCAAGAGATTAACGCGGAGATTTTCGCGCGGAGAAAAGAT 491
Db TTGCGCGGAGAAAAGCAAGAGATTAACGCGGAGATTTTCGCGCGGAGAAAAGAT 244
QY 492 GATCCCGGCAATATTATTCAGTTAGAGTCTCTGTTCTTATGGTGGTCTAAGCGATC 551
Db GATCCCGGCAATATTATTCAGTTAGAGTCTCTGTTCTTATGGTGGTCTAAGCGATC 184
QY 552 TCTGTAAATCTTCGTTAGTTTTCGCGCGTTAAGTCCGCTTCGTTAAGCTGAT 611
Db TCTGTAAATCTTCGTTAGTTTTCGCGCGTTAAGTCCGCTTCGTTAAGCTGAT 124
QY 612 GATTCCCATGGAGTTTATGTTTACCTGCTGCTTCCTTCATGCTTCATTC 671
Db GATTCCCATGGAGTTTATGTTTACCTGCTGCTTCCTTCATGCTTCATTC 64
QY 672 TAATGCTAATACAGATGTAATCTTCAATTTCAATTAATCTGAATTTATTAATGATGA 731
Db TAATGCTAATACAGATGTAATCTTCAATTTCAATTAATCTGAATTTATTAATGATGA 4

RESULT 2
CR356994/c 478 bp DNA linear GSS 05-APR-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-716H04-025111,
DEFINITION genomic survey sequence.
ACCESSION CR356994
VERSION CR356994.1 GI:45539916
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (chale crese)
ORGANISM Arabidopsis thaliana (chale crese)
REFERENCE
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoever, P., and Weisshaar, B.
TITLE GABI-Kat Simplesearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
PUBMED 12874060
REFERENCE
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reis, B., Dekker, K. and Weisshaar, B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics

JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
PUBMED 14756321
REFERENCE
AUTHORS Strizhov, N., Li, Y., Rosso, M.G., Viehoever, P., Dekker, K.A. and Weisshaar, B.

JOURNAL High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
PUBMED Biotechniques 35 (6), 1164-1168 (2003)
REFERENCE 14682050

REFERENCE 4 (bases 1 to 478)
AUTHORS Rosso, M.G., Strizhov, N., Li, Y. and Weisshaar, B.

JOURNAL Direct Submission
TITLE Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g59910. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genome program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES
source
Location/Qualifiers
1..478
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GK-716H04-025111"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pGABI1 (Genbank accession number: AY292716). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 58.7%; Score 430; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.7e-107;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 ACAACCTGTAGATGTTGAGAAATCTCAACTTACAAAAGATTTAATCAATCA 362
Db ACAACCTGTAGATGTTGAGAAATCTCAACTTACAAAAGATTTAATCAATCA 419
QY 363 AGACAGTTATTCGACACCGAGTCATGACGTGCGAAGAAACACCGAAGAAACATGAG 422
Db AGACAGTTATTCGACACCGAGTCATGACGTGCGAAGAAACACCGAAGAAACATGAG 359
QY 418 AGACAGTTATTCGACACCGAGTCATGACGTGCGAAGAAACACCGAAGAAACATGAG 359
QY 423 CTTTCCAGATGTCGCGGAGAAAAGCAAGAGTTACGCGGAGTTCTCGCGCGAGAA 482
Db CTTTCCAGATGTCGCGGAGAAAAGCAAGAGTTACGCGGAGTTCTCGCGCGAGAA 299
QY 358 CTTTCCAGATGTCGCGGAGAAAAGCAAGAGTTACGCGGAGTTCTCGCGCGAGAA 299
QY 483 AAGAGATGATCCGCGGAGATTTATTTCAATTTAGAGTCTGTTCTTATGGTGGTCT 542
Db AAGAGATGATCCGCGGAGATTTATTTCAATTTAGAGTCTGTTCTTATGGTGGTCT 239
QY 298 AAGAGATGATCCGCGGAGATTTATTTCAATTTAGAGTCTGTTCTTATGGTGGTCT 239
QY 543 AACGGATCTGTGTATATCTTCGTTAGTTTTCGCGCGTTACCTTCGCGCTCGTTAT 602
Db AACGGATCTGTGTATATCTTCGTTAGTTTTCGCGCGTTACCTTCGCGCTCGTTAT 179
QY 238 AACGGATCTGTGTATATCTTCGTTAGTTTTCGCGCGTTACCTTCGCGCTCGTTAT 179
QY 603 GCTGCTATGTTGTTCCCATTTGGGATTAATGTTTACCTGCTGCTTCATGATGCTTC 662
Db GCTGCTATGTTGTTCCCATTTGGGATTAATGTTTACCTGCTGCTTCATGATGCTTC 119
QY 178 GCTGCTATGTTGTTCCCATTTGGGATTAATGTTTACCTGCTGCTTCATGATGCTTC 119
QY 663 TTTCATCTCAAGCTAATACAGATGTAATCTTCAATTTCAATTAATCTGAATTTAT 722
Db TTTCATCTCAAGCTAATACAGATGTAATCTTCAATTTCAATTAATCTGAATTTAT 59
QY 723 ATATGATGAT 732
|||||||

| Db | 58 | ATATGATGAT | 49 |
|-----------------------|--|---|----------------------------|
| RESULT 3 | | | |
| BH753057/c | | | |
| LOCUS | | | |
| DEFINITION | BH753057 | 477 bp | DNA linear GSS 27-FEB-2002 |
| | SALK_019707.56.00.x | Arabidopsis thaliana TDNA insertion lines | |
| | Arabidopsis thaliana genomic clone SALK_019707.56.00.x, | genomic | |
| | survey sequence. | | |
| ACCESSION | BH753057 | | |
| VERSION | BH753057.1 | GI:18972604 | |
| KEYWORDS | GSS. | | |
| SOURCE | Arabidopsis thaliana (chale cress) | | |
| ORGANISM | Arabidopsis thaliana | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi. | | |
| REFERENCE | 1 (bases 1 to 477) | | |
| AUTHORS | Alonso,J.M., Leisae,T.J., Barjas,P., Chen,H., Chuk,R., Gaeremab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R. | | |
| | A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome | | |
| | Unpublished (2001) | | |
| JOURNAL | Contact: Joseph R. Ecker | | |
| COMMENT | Salk Institute Genomic Analysis Laboratory (SIGAL) | | |
| | The Salk Institute for Biological Studies | | |
| | 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA | | |
| | Tel: 858 453 4100 x1752 | | |
| | Fax: 858 558 6379 | | |
| | Email: ecker@salk.edu | | |
| | This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 3' end of ATG55900 and 300 bases of the 3' end of ATG55910. | | |
| | Class: TDNA tagged. | | |
| FEATURES | location/Qualifiers | | |
| Source | 1..477 | | |
| | /organism="Arabidopsis thaliana" | | |
| | /mol_type="genomic DNA" | | |
| | /ecotype="Col-0" | | |
| | /db_xref="taxon:3702" | | |
| | /clone="SALK_019707.56.00.x" | | |
| | /clone_lib="Arabidopsis thaliana TDNA insertion lines" | | |
| | /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html " | | |
| ORIGIN | | | |
| Query Match | 54.8%; | Score 401; | DB 9; Length 477; |
| Best Local Similarity | 100.0%; | Pred. No. 1.8e-99; | |
| Matches 401; | Conservative 0; | Mismatches 0; | Indels 0; Gaps 0; |
| 332 | CAAACTTCAAAAAGATATTATTAACATTCACAGAGTTATTCGAACAACGAGTCACAG | 391 | |
| 476 | CAACTTCAAAAAGATATTATTAACATTCACAGAGTTATTCGAACAACGAGTCACAG | 417 | |
| 392 | ACGTCGGAAGAAACAACCGGAAAAACATGAGCTTTCGAAGTTCCGCGAAGAAACAGC | 451 | |
| 416 | ACGTCGGAAGAAACAACCGGAAAAACATGAGCTTTCGAAGTTCCGCGAAGAAACAGC | 357 | |
| 452 | AAGAGTTACGGCGGAGTTTCGCGGCGGAGAAAGAGATGATGCCGCGGAATTATTTCA | 511 | |
| 356 | AAGAGTTACGGCGGAGTTTCGCGGCGGAGAAAGAGATGATGCCGCGGAATTATTTCA | 297 | |
| 512 | GTTTGAAGTCTCTGTTCCATTTGTTGGTCTTACACGCACTCTGTTAATTAATTCGCTAG | 571 | |
| 296 | GTTTGAAGTCTCTGTTCCATTTGTTGGTCTTACACGCACTCTGTTAATTAATTCGCTAG | 237 | |
| 572 | TTTTCGCGCGTTTACCTCCGCGCTCCGTTTATGCTGCTATTGGTTCCATTGGGATTATGG | 631 | |

| DB | TTTTCGCGCGCGTAACTCCGCGCTCCGCTTAAAGCTGCAATTCGTTCCATTCGGATTATATG | 177 |
|-----------------------|--|------------------------|
| QY | 632 TTTTACTCGTCGTCCTTGCCCTTCATCGCCCTTCTTCTCATTCATATGCTAATACAGATGAA | 691 |
| DB | 176 TTTTACTCGTCGTCCTTGCCCTTCATCGCCCTTCTTCTCATTCATATGCTAATACAGATGAA | 117 |
| QY | 692 CTTCGAATTTCAAGTAATCGAAATTTATATATATATGAT | 732 |
| DB | 116 CTTCGAATTTCAAGTAATCGAAATTTATATATATATGAT | 76 |
| RESULT 4 | | |
| BP851800 | BP851800 | 385 bp |
| DEFINITION | BP851800 RAFL21 Arabidopsis thaliana cDNA clone RAFL25-17-007 5', mRNA sequence. | linear EST 17-FEB-2005 |
| ACCESSION | BP851800 | |
| VERSION | BP851800.1 | GI:59912773 |
| KEYWORDS | EST. | |
| SOURCE | Arabidopsis thaliana (thale cress) | |
| ORGANISM | Arabidopsis thaliana | |
| REFERENCE | Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. | |
| AUTHORS | 1 (bases 1 to 385) Seki, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Iida, K., Enju, A., Sakurai, T., Arakawa, T., Carninci, P., Fukuda, S., Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y. and Shinozaki, K. | |
| TITLE | Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs Unpublished (2005) | |
| JOURNAL | Contact: Motoaki Seki | |
| COMMENT | Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel.: 81-298-36-4359 Fax: 81-298-36-9060 Email: meeki@rtc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998, 2002). This clone is in a modified phuescript vector. Please visit our web site (http://pfweb.gsc.riken.jp) and http://range.gsc.riken.jp for further details. location/Qualifiers | |
| FEATURES | 1..385 | |
| SOURCE | /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL25-17-J07" /lab_host="DH10B" /clone_lib="RAFL21" /note="Site 1: BamHI; Site 2: SalI; Subtraction library. The sequence was obtained from samples subjected to various stresses and plant hormones-treated" | |
| ORIGIN | | |
| Query Match | 49.8%; Score 364.6; DB 3; Length 385; | |
| Best Local Similarity | 98.7%; Pred. No. 1.9e-89; | |
| Matches | 378; Conservative 0; Mismatches 4; Indels 1; Gaps 1; | |
| QY | 4 TCTTCCTATTTCCTACTAGTACTTGTTCACAGATTCTTGATCAACCAACCA | 63 |
| DB | 1 TCTTCCTATTTCCTACTAGTACTTGTTCACAGATTCTTGATCAACCAACCA | 60 |
| QY | 64 TACCAAAAGCTTCGAAATCTTCACCTCAAGCTTCTTCCTTAACTCGAATGCTG | 123 |
| DB | 61 TACCAAAAGCTTCGAAATCTTCACCTCAAGCTTCTTCCTTAACTCGAATGCTG | 120 |
| QY | 124 AGTAACTCGGATTGTTTCGACATCCCTGTTTTCGAATGCTGGGCATCTTATTTGT | 183 |
| DB | 121 AGTAACTCGGATTGTTTCGACATCCCTGTTTTCGAATGCTGGGCAT-CTTATTTGT | 179 |
| QY | 184 CTGCAATTCCTGACCAATTCGCTGATCAAGCTGATTCGTTAACAGTTGCCCTAAGA | 243 |

```

Db      180 CTGAATCTTCACCAATGCTTCATCAAGCTCAGTGGTTACCAAGTTCCTTAAGA 239
Qy      244 TCGATCTTTGAGCAAAATTTTGTCACTGATCTTCTTAAATCCAAACGACGACGAAA 303
Db      240 TCGATCTTTGAGCAAAATTTTGTCACTGATCTTCTTAAATCCAAACGACGACGAAA 299
Qy      304 CAACCTCTGATAGATTCGAGAAATCTCAAACTTACAAAAAGATTTATTAACATTCAA 363
Db      300 CAACCTCTGATAGATTCGAGAAATCTCAAACTTACAAAAAGATTTATTAACATTCAA 359
Qy      364 GACAGTTATTCGAACCAACCGAGT 386
Db      360 GACAGTTATTCGAACCAACCGAGT 382

RESULT 5
LOCUS   CL485455                      920 bp  DNA          linear  GSS 01-Apr-2004
DEFINITION
SAIL_416_D03.v1 SAIL Collection Arabidopsis thaliana genomic clone
ACCESSION
CL485455
VERSION
CL485455.1 GI:45962135
KEYWORDS
GSS.
SOURCE  Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 920)
Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
Dietrich,B., Ho,P., Baccaden,J., Ko,C., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Miguel,T., Hutchinson,D., Kimerly,B.,
Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
12468722

REFERENCE
AUTHORS  Contact: Sessions A
JOURNAL  Applied Trait Genetics
PUBMED   Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS819225; T-DNA left border flanking sequences of
Syngenta Arabidopsis insertion library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1..920
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_416_D03.v1"
/clone_1kb="SAIL Collection"
/notes="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"

ORIGIN

```

```

Query Match      35.8%; Score 262.4; DB 10; Length 920;
Best Local Similarity 70.5%; Pred. No. 4e-61;
Matches 437; Conservative 0; Mismatches 171; Indels 12; Gaps 6;

Qy      45 TGCATCCACCAAAACCAATACACAAAGCTTTCAACTCTTCACCTCAAGCTTCTTC 104
Db      250 TGCATCAAGCAACTCCCGTACAAAGACGCTGTGGAATCTTCATGCCCCAGCTACGCA 309
Qy      105 CTTTACATCGAATCGTTGATTAAGTAACTCGAT-TTGTTTGCATACCTCGTTTCTGAATC 163
Db      310 CTTTACGCGGAAAACCTACCTACCACTTCTACGAGGCTGAGACCTTAAGCACTCAATC 369
Qy      164 GTGGGCGATCCTTATTTGTCTGCAATCTTCAACCAATTCGATCGATCAAGCTGCAATGG 223

```

```

Db      370 GAGAGGCATCTTATTATTAACCGTCAATATTATACCAATTAAGGAGCTAAACCACTTCAG 429
Qy      224 TTAACCAAGTGGCCCTTAAGATCAAGATTTTGACAAAATTTGTCACTGATCTTGAAT 283
Db      430 ATACCCAG-----ACAGCTTATTAACCTGCACATATGGAGAGTGTTCCTCCG 482
Qy      284 CCAACGACACAGCAAAACCAACCTCTGATAGTANT-CGAGAAATCTGAACCTTCAA 342
Db      483 TCAATCCACACACAAACCAACCTCTCGATGATGATACGAGATTTCTAAACTTACAA 542
Qy      343 AAAAATATTAATAACATTCAGAAGTATTCGAACCAACCAAGTATGACGTGCAAGA 402
Db      543 AAGATCTTATTAACATTCAGAAGTATTCGAACCAACCAAGTATGACGTGCAAGA 602
Qy      403 AACACCGGAAAAACATGAGCTTTCGAAGTTCGCGGAGAAAAGCAAGCAAGTTACGG 462
Db      603 AACCAACC-GTAAACACATGAGCTTTCGAAGTTCGCGGAGAAAAGCAAGCAAGTTACGG 661
Qy      463 CGGAGTTTCTCGGCGGCAAAAGATGATATCCCGCGAATTAATTACGTTAGAGTCT 522
Db      662 CGGAGTTTCTCGGCGGCA-AAAAAGATGATATCCCGCGAA-TATTTCAAGTTAGAGTCT 719
Qy      523 CTGTTCTATTGTTGTTGTTTACGCGATCTCTGTTAATACCTTCGTTAGTTTGC CGG 582
Db      720 CTGTCCTATCAGAGGGGGCTTAACGCAATCTCTGTTAATACCTTCGTTAGTTTGC CGG 779
Qy      583 TTACCTTCGCTCCGTTTATGCTGCTATGTTGTTCCATTCGGAATATGTTTATCTGTC 642
Db      780 TTACCTTACGCTCCGTTTATGCTGCTAAGGTTCCACCTGGAGCTATGCGTAACCCCG 839
Qy      643 GTTCTTGCTTCATGCTTC 662
Db      840 GGAATGGGCTTAAACCTTAC 859

```

```

RESULT 6
LOCUS   CC963171/c                      639 bp  DNA          linear  GSS 18-Aug-2003
DEFINITION
BOEM18TF_BO_1.4_1.6_KB_nuc Brassica oleracea genomic clone
BOEM18, genomic survey sequence.
ACCESSION
CC963171
CC963171.1 GI:33812283
VERSION
CC963171.1
KEYWORDS
GSS.
SOURCE  Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 639)
Ayala,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utechtack,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
Other_GSSs: BOEM18TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..639
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOEM18"

FEATURES
source

```

```


```


ORIGIN

/clone_11b="BO 1.4 1.6 KB nuc"
/note="Vector: pHO52; Site 1: BstXI; 1.4-1.6 kb shared
nuclear DNA inserted into pHO52 using BstXI linkers"

Query Match 32.7%; Score 239; DB 9; Length 639;
Best Local Similarity 80.1%; Pred. No. 1.1e-54;
Matches 326; Conservative 0; Mismatches 60; Indels 21; Gaps 3;

ORIGIN

/ecotype="Col-0"
/db_xref="taxon:3702"
/clone_11b="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"

Query Match 26.1%; Score 191; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

316 ATGATTGAGAAATCTCAACTTACAAAAGATTATTAACATTCAAGACATTATTG 375
404 ATGATTGAGAAACCCCAATTTACAGAACATCATTAACATCAAGACGTTACTCA 345
376 AACACCGAGTCATGAC-----GTGGAGAGAAACACCGGAAACATAGCTTCA 429
344 AACACCTTGTCTGACGACGAGAGAGAGAAACACCGGAAACATTAACCTTTCG 285
430 AGTTGCGCG-----GAGAAAGCAAGCAAGATTACGGCGAGTTCTCGCG 477
284 CGTCACCGGCAATGAGGTGAGAGAACAGAGATGAGTTGCGACGACCTTCTCG 225
478 CAGAAAGAGATGATCCCGCAATTATTTAGATTAGAGTCTGTTCTTAATGGTT 537
224 CAGAAAGATTGATCCCGCAACTACTGATTGAGTCTGTTATCTGGTG 165
538 GGTCTAACGGCATCTGTTAATCTCCGTTAGTTTCCGCGTTACCTCCGCGCTCG 597
164 GGTCTAACGGCGCATCTGTTATCTACGTTAGTTTCCGCGTTACCTCCGCGCTCG 105
598 TTATGCTGCTATGTTCCCATTTGAGTTATGTTTACTGTTCTTCCCTTCAAG 657
104 TTATGCTGCTTGTGATTCCCATTTGGGATTAATGTTTACTGTTCTTCCCTTCAAG 45
658 CTTTCTTCTCATTTTAATGCTAATACAGATGTAATCTGCAATTTCAAT 704
44 CTTTCTTCTTCTTCTAAGACCA--AGATGTAATCACTTAT 1

RESULT 7 321 bp DNA linear GSS 04-SEP-2002
BH904895 SALK_105296.54.75.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_105296.54.75.x, genomic
survey sequence.

ACCESSION 321 bp DNA linear GSS 04-SEP-2002
BH904895 SALK_105296.54.75.x Arabidopsis thaliana TDNA insertion lines
VERSION BH904895.1 GI:22717538
KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 321)

REFERENCE Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrihab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shim,P., Zimmermann,J. and Becker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)

JOURNAL COMMENT
TITLE The Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.

FEATURES
source

Class: TDNA tagged.
Location/Qualifiers
1..321
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"

61 CAATACAAAGCTTCTCAAACTCTTCACTCAAGCTTCTTCTTACATCTGAATCG 120
191 CAATACAAAGCTTCTCAAACTCTTCACTCAAGCTTCTTCTTACATCTGAATCG 250
121 TTGAGTAACTCGGATTGTTCTGCAATCTCTGTTCTGAATCGTGGGCAATCTTATT 180
251 TTGAGTAACTCGGATTGTTCTGCAATCTCTGTTCTGAATCGTGGGCAATCTTATT 310
181 TGCTCGAATT 191
311 TGCTCGAATT 321

RESULT 8 371 bp DNA linear GSS 04-SEP-2002
BH904894 SALK_105295.56.00.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_105295.56.00.x, genomic
survey sequence.

ACCESSION 371 bp DNA linear GSS 04-SEP-2002
BH904894 SALK_105295.56.00.x Arabidopsis thaliana TDNA insertion lines
VERSION BH904894.1 GI:22717536
KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 371)

REFERENCE Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrihab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shim,P., Zimmermann,J. and Becker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)

JOURNAL COMMENT
TITLE The Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.

FEATURES
source

Class: TDNA tagged.
Location/Qualifiers
1..371
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone_11b="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines"

each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html

ORIGIN

Query Match 26.1%; Score 191; DB 9; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCTTCTCATTTCCCTACTAGTACTGTTTTCACACAGTTCTTGATCCAAACAAAC 60
DB 181 TTGCTTCTCATTTCCCTACTAGTACTGTTTTCACACAGTTCTTGATCCAAACAAAC 240
QY 61 CAATPACAAAGCTTCCAAAGCTCTCACTCAAGCTTCCCTTACATCGAATCG 120
DB 241 CAATPACAAAGCTTCCAAAGCTCTCACTCAAGCTTCCCTTACATCGAATCG 300
QY 121 TTGAGTTAATCGAATTTGTTGTCATCTCTGTTTGAATCGGGCCATCCTTAATT 180
DB 301 TTGAGTTAATCGAATTTGTTGTCATCTCTGTTTGAATCGGGCCATCCTTAATT 360
QY 181 TGCTCGAATT 191
DB 361 TGCTCGAATT 371

RESULT 9

AVS44303/c 471 bp mRNA linear EST 20-FEB-2004
LOCUS AVS44303 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION cDNA clone R239a09f 3', mRNA sequence.
ACCESSION AVS44303
VERSION AVS44303.1 GI:8715717
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 471)
Aamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries

JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847
COMMENT Contact: Erika Aamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1533-3, Kisarazu, Chiba 292-0812, Japan
Email: aamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

FEATURES
SOURCE
1. 471
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="R239a09f"
/tissue="roots"
/clone_id="Arabidopsis thaliana roots Columbia"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN

Query Match 24.9%; Score 182.6; DB 1; Length 471;
Best Local Similarity 70.0%; Pred. No. 3.8e-39;
Matches 296; Conservative 0; Mismatches 109; Indels 18; Gaps 3;

QY 270 CTGATCTTCTTAATCCAAACGACACAGCAAACTCTGTAGATGATTCGAAAT 329
DB 460 CTGATTTCCAGATCAAACTCGAAAAGAAAAGCCTTCTTAATGATTCGTGAGTT 401

QY 330 CTCGAACCTTCAAAAAGATTTATTAACATTCAAGACAGTATTGCAACACGAGCAT 389
DB 400 CTCGAGCTTCAAAAAGACATCATTAACATTTCAAGAACATTTCTCTCAACAAACAT 341
QY 390 GGAAGTCGAGAGA---AACACCGGAAAAACATGAGCTTTCGAATT-----CGCC 437
DB 340 GGAAGTCGAGAGAGATCATTAACCGGAAAAACAGAGTTTTCGAGTTTCAAGCTCAGCTC 281
QY 438 GGAAGAACGACACAGAGATTAACGGAGATTCTCGGCGCAGAA-----AGAGATAT 491
DB 280 GATTATGGGAGAACAGAAATTTGTTTCGACATTTGTCGACAAACAGTCAAGAGAGCT 221
QY 492 GATCCGGGCAATTAATTTGATTAGAGTCTGTTCTTAATGATGATGATGATGATGATG 551
DB 220 AATATACGAGATTAATTTGATTAGATTAATGATGATGATGATGATGATGATGATG 161
QY 552 TCTGTTAATACCTTCGTTAGTTTTCGCGCTTACCTCGCTCGTTAATGCTGAT 611
DB 160 TCTGTTAATACCTTCGTTAGTTTTCGCGCTTACCTCGCTCGTTAATGCTGAT 101
QY 612 GATTCGATTTGGGATTAATGATTTTACTGTCGTTCTTCTGCTTATGCTTCTCATTC 671
DB 100 GATTCGATTTGGGATTAATGATTTGCTTATGTTTCTTTCATGCTTCTTCAATTC 41
QY 672 TAA 674
DB 40 CAA 38

RESULT 10

CNS0A8DR 763 bp mRNA linear HTC 06-FEB-2004
LOCUS CNS0A8DR
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSUTS1L73ZF02 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION BX821785
VERSION BX821785.1 GI:42467792
KEYWORDS HTC; GSUT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 763)
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Queller, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

FEATURES

FEATURES
SOURCE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 763)
Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequenage ;
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_FR/Full
length

FEATURES
SOURCE
1. 763
Location/Qualifiers
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jailion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. UMG INRA : Clepet C., Caboche M.
FEATURES Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_FR/Full length
source Location/Qualifiers
 1. 850
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSLRPH94ZB08"
 /tissue_type="Hormone Treated Callus"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"
 1. 850
 /gene="At2g44080"
ORIGIN
 Query Match 16.1%; Score 118; DB 4; Length 850;
 Best Local Similarity 61.7%; Pred. No. 3.1e-21;
 Matches 264; Conservative 0; Mismatches 145; Indels 19; Gaps 4;
 QY 270 CTGATCTTAAATCCAAACGACACGAAACAACTCTGTAGATGATTCGAGAAAT 329
 DB 377 CTGATTTTCAGATCAAACTTCGAAACAGAAAGAGCTTTAAATCATTCGTCAATT 436
 QY 330 CTCGAACCTTCAAAAAGATTTAATTAACATTCGAAAGACATTCGAAACCGAGATCAT 389
 DB 437 CTCGAGCTCAAAAACCAATCATACATTTGAAGAAACATATCTCTCCAAACAACAT 496
 QY 390 GGAAGTCGGAAGAAC---AACGGAAAAACATGAGCTTCGAAAGTTCG-----CC 437
 DB 497 GCACTCAGAGCGGATATACCGGACACACAGAGTTCGTGTTCACTCCAGCTCC 556
 QY 438 GGAAGAAAGCAAGACAGATTACGGCGGAGTTCTCGCGCAGAA-----AAGATGAT 491
 DB 557 GATTATGGGAGAGGAGAGTTGTTTCGACATTTGTGGGAGAGAGTCCAAGAGAGCT 616
 QY 492 GATCCCGCGAATTAATTCGTTAGAGTCTCT-GTTCCTATTTGTTGTTCTAACGGCAT 550
 DB 617 CATATCAGCAGATCACTTCAGTTAGATGAATGATGTTGTTGGGTTCTGAAAAAT 676
 QY 551 CTCGTAAATACCTTCGTAGTTTGGCGCGTTACTCCGCTCGTTATGCTGCTAT 610
 DB 677 ATCTCTGATCTTACCGTTGATCTTCCACCATTCCTCTCTCTTTATGCTATTT 736
 QY 611 TGGTCCCATTTGGATTATGATTACTGTCGTCTTTCGCTTCATGCTTCATTCAT 670
 DB 737 TGAATTCCTATGGGATATATGTTTGTATGTTGGGAGTTGATGCGGGGCGGATTT 796
 QY 671 CTAAATGCT 678
 DB 797 TCAAAAT 804

RESULT 13
 BZ039389 793 bp DNA linear GSS 09-OCT-2002
 LOCUS Ikaz0106.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
 DEFINITION sequence.

ACCESSION BZ039389
VERSION BZ039389.1 GI:23626714
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Deleantony K., Fewell G., Fulton L., McComble W.R., Miner T., Nash W., Radlowicz P.D. and Wilson R.K. Whole genome shotgun reads from Brassica oleracea Unpublished (2002)
AUTHORS Contact: Richard K. Wilson
TITLE Genome Sequencing Center
JOURNAL Washington University School of Medicine
COMMENT Email: submissions@watson.wustl.edu
 Plate: Ikaz0 row: f column: 06
 Seq primer: -28RPOT reverse
 Class: shotgun
 High quality sequence start: 77
 High quality sequence stop: 511.
FEATURES Location/Qualifiers
 1. 793
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone="B.oleracea002"
 /note="Vector: pOTW13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Radlowicz (GSH) and the shotgun library prepared at Washington University Genome Sequencing Center."
ORIGIN
 Query Match 15.7%; Score 114.8; DB 9; Length 793;
 Best Local Similarity 72.3%; Pred. No. 2.3e-20;
 Matches 149; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
 QY 480 GAAAGATGATGATCCGCGAATTAATTCAGTTAGAGTCTCTGCTATGTTGG 539
 DB 513 GAGAGGCTAAATTAATTAATTCAGTTAGAGTCAATGATGATGTTGTTGG 454
 QY 540 TCTAAGGCAATCTCTGTTAATTAATTCGTTAGTTTGGCGGCTTACCTCGCTCCGTT 539
 DB 453 TCTCAGAGATGCTCTCTGTTCCCTTCGTTGATTCCTCAACCTTCCCTCCCTT 394
 QY 600 TATGCTGATATGTTCCCATTTGAGATTATGTTTACTCGTGTCTTGCCTTCAAGCC 659
 DB 393 CATGCTGCTTCTGTTCTCTATTTGAGATTATGTTTGTCTTATGTTCTTCTTAAATGCC 334
 QY 660 TTCTTCTCATTTCAATGCTAATACAG 685
 DB 333 TTCCTTAATGCAAAACATGTAACAG 308

RESULT 14
 CV432287 474 bp mRNA linear EST 28-SEP-2004
 LOCUS RT0425 Chinese cabbage root library Brassica rapa subsp. pekinensis
 DEFINITION CDNA clone RT0425 5', mRNA sequence.
ACCESSION CV432287
VERSION CV432287.1 GI:52827409
KEYWORDS EST.
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustoids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 474)
 Ryu, S.H., Kang, J.S., Yang, K.A. and Lim, C.O.

| TITLE | Expressed Sequence Tags of Chinese cabbage roots |
|---------------------------|--|
| JOURNAL | Unpublished (2004) |
| COMMENT | Contact: Lim, C.O. Plant Molecular Biology & Biotechnology Research Centre Gyeongsang National University #900 Gazwa-dong, Jinju 660-701, Korea Tel.: 82 55 751 6255 Fax: 82 55 759 9363 Email: colim@nongae.gsnu.ac.kr Seq primer: M13 universal primer. |
| FEATURES | Location/Qualifiers |
| SOURCE | 1..474 |
| ORIGIN | /organism="Brassica rapa subsp. pekinensis" /mol_type="mRNA" /cultivar="Dangwon" /db_species="pekinensis" /db_xref="taxon:51351" /clone="RT0425" /dev_stage="Seedling" /lab_host="XL-1 Blue" /clone_1fb="Chinese cabbage root library" /note="Organ: Root; Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I" |
| Query Match | 14.0%; Score 102.8; DB 7; Length 474; |
| Best Local Similarity | 57.8%; Pred. No. 4.3e-17; |
| Matches 245; Conservative | 0; Mismatches 167; Indels 12; Gaps 3 |
| OY | 290 CAGACAGCGCAAAACAACCTCTGTGATGATTCCGGAATTCCTCAACTTACAAAAGATA 349 |
| Db | 49 CAGAAAAAAGACTACACCCTTTTATGATGATTCGGAATTCCTGTCTACAAAACGACA 108 |
| OY | 350 TTATTAACATTCAGAAGCAGATTATTCGAACAACCGATCATGGACCTCGGAAGAACCAAC 409 |
| Db | 109 TCGTAAACATTCAGAAGCAGCTCTTCAACAGCAACAAGATCATGACATGAGAGAG 168 |
| OY | 410 GGAAAAACATGAGCTTGGAAGTTGCCGAGAAAAACAAGACAGTAGTTACGCGGAGTT 469 |
| Db | 169 ATATTCGCGCACACAGCTCAGCT---CCTAATGATGATGAACAGAAATATTTTCGACAT 225 |
| OY | 470 TCTGGGCGGCAGAAA-----AGGATGATGATCCCGGCAATATTTACGTTAGAGTCTC 523 |
| Db | 226 TGTCGCTCAAGAACAGTCCGAGAGAGCTPAATCTCAGCAGATTACTTCAGTTGAGTCAA 285 |
| OY | 524 TGTTCCTATTTGATGATCTTAACGCGCATCTGTATATCTTCGTTAGTTTGCCGCGCT 583 |
| Db | 286 TGTGTTCCTTGTGTGTCTCAGACAGATGCTCTTATCTTCTTGATTTCTT---CACCG 342 |
| OY | 584 TAACCTCGCCCTCCGTTTATGCTGATTTGATTCGATTCGATTTATGATTTTACTCGTCG 643 |
| Db | 343 TTGCTCCCTCTCCTTTTCAATGCTGCTTCGATTCGATTCGATTCGATTCGATTCGATTCG 402 |
| OY | 644 TTCCTGCTTATGCTGCTTCTTCTATCTTATGCTTATTCAGATGCTTATTCGATTTCA 703 |
| Db | 403 TCTGCTTATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 462 |
| OY | 704 TGTG 707 |
| Db | 463 TGTG 466 |
| RESULT 15 | |
| LOCUS | CC737187/c |
| DEFINITION | CC737187 765 bp DNA linear GSS 27-JUN-2003 |
| ACCESSION | A1164_Maxxa-2_A09 Maxxa Gooseyplum hitsutum genomic, genomic survey |
| VERSION | CC737187 |
| KEYWORDS | CC737187.1 GI:32306703 |
| SOURCE | GSS. |
| ORGANISM | Gooseyplum hitsutum (upland cotton) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; |

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|-----------|---|--|----------------------------------|---------|
| 1 | Rong, J., Abbey, C., Bowers, J.E., Brubaker, C.L., Chang, C., Chee, P.W., Delmonte, T.A., Ding, X., Garza, J.J., Martler, B.S., Park, C.H., Pierce, G.J., Rainey, K.M., Rastogi, V.K., Schuler, S.R., Trolander, N.L., Wendel, J.F., Wilkins, T.A., Williams-Coplin, T.D., Wright, R.A., Wright, R.J., Zhao, X., Zhu, L. and Paterson, A.H. | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Malvales; Malvaceae; Malvaceae; Gossypium. | Genetics 166 (1), 389-417 (2004) | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |
| 5 | | | | |
| 6 | | | | |
| 7 | | | | |
| 8 | | | | |
| 9 | | | | |
| 10 | | | | |
| 11 | | | | |
| 12 | | | | |
| 13 | | | | |
| 14 | | | | |
| 15 | | | | |
| 16 | | | | |
| 17 | | | | |
| 18 | | | | |
| 19 | | | | |
| 20 | | | | |
| 21 | | | | |
| 22 | | | | |
| 23 | | | | |
| 24 | | | | |
| 25 | | | | |
| 26 | | | | |
| 27 | | | | |
| 28 | | | | |
| 29 | | | | |
| 30 | | | | |
| 31 | | | | |
| 32 | | | | |
| 33 | | | | |
| 34 | | | | |
| 35 | | | | |
| 36 | | | | |
| 37 | | | | |
| 38 | | | | |
| 39 | | | | |
| 40 | | | | |
| 41 | | | | |
| 42 | | | | |
| 43 | | | | |
| 44 | | | | |
| 45 | | | | |
| 46 | | | | |
| 47 | | | | |
| 48 | | | | |
| 49 | | | | |
| 50 | | | | |
| 51 | | | | |
| 52 | | | | |
| 53 | | | | |
| 54 | | | | |
| 55 | | | | |
| 56 | | | | |
| 57 | | | | |
| 58 | | | | |
| 59 | | | | |
| 60 | | | | |
| 61 | | | | |
| 62 | | | | |
| 63 | | | | |
| 64 | | | | |
| 65 | | | | |
| 66 | | | | |
| 67 | | | | |
| 68 | | | | |
| 69 | | | | |
| 70 | | | | |
| 71 | | | | |
| 72 | | | | |
| 73 | | | | |
| 74 | | | | |
| 75 | | | | |
| 76 | | | | |
| 77 | | | | |
| 78 | | | | |
| 79 | | | | |
| 80 | | | | |
| 81 | | | | |
| 82 | | | | |
| 83 | | | | |
| 84 | | | | |
| 85 | | | | |
| 86 | | | | |
| 87 | | | | |
| 88 | | | | |
| 89 | | | | |
| 90 | | | | |
| 91 | | | | |
| 92 | | | | |
| 93 | | | | |
| 94 | | | | |
| 95 | | | | |
| 96 | | | | |
| 97 | | | | |
| 98 | | | | |
| 99 | | | | |
| 100 | | | | |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 30, 2006, 08:59:59 ; Search time 3791 Seconds

(without alignments)
1308.212 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533
Sequence: 1 MDVGRNKRKMSFRSSPEKS.....LAFMPSHSNANDVTGNCNM 106

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPRO.spool.p/US10715129/runat.27012006.154745.14579/app.query.fasta.1.263
-DB=EST -ORMT=fagcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biolum62 -TRANS=human40.cdi -LIST=45
-DOCLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10715129@cgn.1.1.8010 @runat.27012006.154745.14579 -NCPU=6 -ICPU=3
-NO MMAP -MBG SCORES=0 -WAIT -DSPBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7 -YGAPOP=10
-YGAPEXT=0.5 -DELop=6 -DELext=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 533 | 100.0 | 477 | 9 | BH753057 SALK_0197 |
| 2 | 533 | 100.0 | 478 | 11 | CR356994 Arabidops |
| 3 | 518 | 97.2 | 543 | 1 | AI998680 701546552 |
| 4 | 395.5 | 74.2 | 639 | 9 | CC963171 BOEIM187F |
| 5 | 332 | 63.3 | 763 | 4 | CNS0A8DR |
| 6 | 327 | 61.4 | 471 | 1 | AV544303 |
| 7 | 292 | 54.8 | 616 | 9 | BH528419 BOGCC46TR |

| | | | | | | |
|---|----|-------|------|-----|----|----------|
| c | 8 | 274 | 51.4 | 793 | 9 | BZ039389 |
| c | 9 | 270.5 | 50.8 | 382 | 2 | BG125370 |
| c | 10 | 270.5 | 50.8 | 624 | 1 | AI774560 |
| c | 11 | 259.5 | 48.7 | 678 | 5 | BW688619 |
| c | 12 | 258.5 | 48.5 | 606 | 3 | BI934590 |
| c | 13 | 257.5 | 48.3 | 635 | 7 | BI932375 |
| c | 14 | 257.5 | 48.3 | 657 | 7 | CN192376 |
| c | 15 | 257.5 | 48.3 | 749 | 8 | DN625498 |
| c | 16 | 257.5 | 48.3 | 752 | 7 | CP831325 |
| c | 17 | 257.5 | 48.3 | 775 | 7 | CN187913 |
| c | 18 | 257.5 | 48.3 | 779 | 7 | CN188780 |
| c | 19 | 257.5 | 48.3 | 859 | 8 | CX045872 |
| c | 20 | 256.5 | 48.1 | 678 | 8 | DN587342 |
| c | 21 | 256.5 | 48.1 | 741 | 7 | CP829407 |
| c | 22 | 256 | 48.0 | 578 | 6 | CP404276 |
| c | 23 | 255.5 | 47.9 | 661 | 7 | CV476718 |
| c | 24 | 255.5 | 47.9 | 694 | 7 | CV255174 |
| c | 25 | 255.5 | 47.9 | 817 | 7 | CV243132 |
| c | 26 | 254.5 | 47.7 | 859 | 8 | CX668291 |
| c | 27 | 254.5 | 47.7 | 944 | 8 | CX668915 |
| c | 28 | 254.5 | 47.7 | 973 | 8 | CX664941 |
| c | 29 | 254 | 47.7 | 758 | 8 | CX655014 |
| c | 30 | 254 | 47.7 | 771 | 7 | CV266016 |
| c | 31 | 253.5 | 47.6 | 369 | 7 | CN548499 |
| c | 32 | 253.5 | 47.6 | 683 | 7 | CF830612 |
| c | 33 | 253.5 | 47.6 | 688 | 7 | CF831943 |
| c | 34 | 253.5 | 47.6 | 698 | 7 | CF829409 |
| c | 35 | 252.5 | 47.4 | 900 | 7 | CF829408 |
| c | 36 | 250.5 | 47.0 | 704 | 7 | CF830612 |
| c | 37 | 239.5 | 44.9 | 622 | 7 | CK093413 |
| c | 38 | 238.5 | 44.7 | 610 | 3 | BP929940 |
| c | 39 | 237.5 | 44.6 | 551 | 7 | CV237738 |
| c | 40 | 237.5 | 44.6 | 598 | 8 | CX175107 |
| c | 41 | 237.5 | 44.6 | 706 | 8 | CX174683 |
| c | 42 | 237.5 | 44.6 | 794 | 7 | CV257043 |
| c | 43 | 237 | 44.5 | 577 | 8 | DN625497 |
| c | 44 | 236.5 | 44.4 | 920 | 10 | CI485455 |
| c | 45 | 234 | 43.9 | 581 | 2 | BG317360 |

ALIGNMENTS

RESULT 1
BH753057/c 477 bp DNA linear GSS 27-FEB-2002
LOCUS SALK_019707.56.00.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_019707.56.00.x, genomic survey sequence.

ACCESSION BH753057
VERSION BH753057.1 GI:18972604

KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gnatelab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shin, P., Zimmerman, J., and Becker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)

TITLE JOURNAL
COMMENT The Salk Institute Genomic Analysis Laboratory (SIGAL)
Contact: Joseph R. Becker
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckers@salik.edu

This is single pass sequence recovered from the left border of TMDA. This sequence lies within 300 bases of the 3' end of AT395990 and 300 bases of the 3' end of AT3959910.

Class: TDNA tagged.

FEATURES
source

Location/Qualifiers
1..477
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_019707.56..00.x"
/note="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Alignment Scores:

Pred. No.: 8.17e-48 Length: 477
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-715-129-2 (1-106) x BH753057 (1-477)

QY 1 MetAapValGlyArgAnaAnaArgLySaMserPheAArgSerProGluLysSer 20
DB 420 ATGAGCGTCGGAAACAAACCGAAGAAACATGAGTTTCGAAGTTCCGCGAGAAAGC 361
QY 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLysArgMetCileProAlaAsnTyr 40
DB 360 AAGCAAGATTACCGCGAGTTTCTCGCGCGAGAAAGATGATGATCCGCGAATTAT 301
QY 41 PheSerLeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
DB 300 TTCAGTTTAAAGTCTCGTTCTTATGTTGTTCTAACGCGATCTCTGTAATACTTCG 241
QY 61 LeuValLeuProProLeuProProProPheMetLeuLeuLeuValProIleGlyIle 80
DB 240 TTAGTTTGGCGCGCTTACCTCCGCTCCGTTATGCTGATGTTCCCATTTGGGATT 181
QY 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
DB 180 ATGCTTTTACCTCGCGTTCTTGGCTTATGCTTCTTCTATCTATATGTAATACGAT 121
QY 101 ValThrCysAsnPheMet 106
DB 120 GTAACCTGCATTTTCATG 103

RESULT 2

CR356994/c 478 bp DNA linear GSS 05-Apr-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-716H04-025111,
DEFINITION genomic survey sequence.
ACCESSION CR356994
VERSION CR356994.1 GI:45539916
KEYWORDS GSS.

SOURCE
ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
AUTHORS Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weisshaar, B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)

REFERENCE
PUBMED 12874060
AUTHORS Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and

TITLE

Weisshaar, B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
PUBMED 14756321
REFERENCE

AUTHORS

Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
Weisshaar, B.
High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
Biotechniques 35 (6), 1164-1168 (2003)
JOURNAL 14682050

PUBMED

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany.
This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene At3G59910.
Details on the protocols used for generation of the sequence are
described in References 1-3. The sequences are generated at the MPI
for Plant Breeding Research in the context of the GABI-Kat project.
GABI-Kat is part of the German Plant Genomics program designated
'GABI'. Information on line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers

1..478
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GK-716H04-025111"
/clone_1b="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pGABI1 (Genbank accession number: AY529716). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."

ORIGIN

Alignment Scores:

Pred. No.: 8.19e-48 Length: 478
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-715-129-2 (1-106) x CR356994 (1-478)

QY 1 MetAapValGlyArgAnaAnaArgLySaMserPheAArgSerProGluLysSer 20
DB 393 ATGAGCGTCGGAAACAAACCGAAGAAACATGAGTTTGAAGTTCCGCGAGAAAGC 334
QY 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLysArgMetCileProAlaAsnTyr 40
DB 333 AAGCAAGATTACCGCGAGTTTCTCGCGCGAGAAAGATGATGATCCGCGAATTAT 274
QY 41 PheSerLeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
DB 273 TTCAGTTTAAAGTCTCGTTCTTATGTTGTTCTAACGCGATCTCTGTAATACTTCG 214
QY 61 LeuValLeuProProLeuProProProPheMetLeuLeuLeuValProIleGlyIle 80
DB 213 TTAGTTTGGCGCGCTTACCTCCGCTCCGTTATGCTGATGTTCCCATTTGGGATT 154
QY 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
DB 153 ATGCTTTTACCTCGCGTTCTTGGCTTATGCTTCTTCTATCTATATGTAATACGAT 94
QY 101 ValThrCysAsnPheMet 106

Db 93 GTAACCTGCATTCATG 76

RESULT 3
LOCUS A1998680/c 543 bp mRNA linear EST 08-SEP-1999
DEFINITION 701546552 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis thaliana cDNA clone 701546552, mRNA sequence.

ACCESSION A1998680
VERSION A1998680.1 GI:5845585
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Chen, J., Momiyama, M., Chan, B., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoka, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kasuri, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nodriguez, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.
TITLE Arabidopsis thaliana Gene Expression Microarray
JOURNAL Unpublished (1999)
COMMENT Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES
SOURCE
1. 543
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="701546552"
/tissue_type="rosette"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-2"
/note="Vector: pSPORT, Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was bluntended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN
Alignment Scores:
Pred. No.: 4.1e-46 Length: 543
Score: 518.00 Matches: 104
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 2
Query Match: 97.2% Indels: 0
Gaps: 0
DB: 1
US-10-715-129-2 (1-106) x A1998680 (1-543)

Qy 1 MetApValGlyArgAsnAraGlyAsnMetSerPheArgSerSerProGluLysSer 20
Db 347 ATGACCTCGAAGAAACACCGAAGAAACATGACCTTCAGATTGCGCGAAGAAAGC 288

Qy 21 LygGInGluLeuAraGArgSerPheSerAlaGlnLysArgMetCilieProAlaAsnTyr 40
Db 287 AAGCAAGATTACGCGCGAGATTCTCGCGCGCAAGAAAGATGATGATCCCGCGCAATTAT 228

Qy 41 PheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
Db 227 TTCAGTTAAGATCTGCTTCTAATGATGCTTAACGGATCTCTGTAATCTCCG 168

Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuValProIleGlyIle 80
Db 167 TTACGTTTGGCGCGCGTTACNMTCCGCCCTCGTTATGCTGCAATGGTTCCCATGGGATT 108

Qy 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
Db 107 ATGGTTTACTCGCGCTTGTTCCTTCATGCTTCATCTCAATGCTAATACAGAT 48

Qy 101 ValThrCysAsnPhenMet 106
Db 47 GTTAACCTGCATTCATG 30

RESULT 4
LOCUS CC963171/c 639 bp DNA linear GSS 18-AUG-2003
DEFINITION BOEM18TF BO_1.4_1.6_KB_nuc Brassica oleracea genomic clone
BOEM18, genomic survey sequence.

ACCESSION CC963171
VERSION CC963171.1 GI:33812283
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE
AUTHORS Ayale, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utezhack, T.R., Wortman, J.R., White, O.R. and Town, C.D.
TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL Genome Res. 15 (4), 487-495 (2005)
PUBMED 15805490
COMMENT Other_GSSs: BOEM18TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.
location/Qualifiers
1. 639
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOEM18"
/clone_lib="BO_1.4_1.6_KB_nuc"
/note="Vector: pHO2; Site_1: BseXI; 1.4-1.6 kb sheared nuclear DNA inserted into pHO2 using BseXI linkers"

ORIGIN
Alignment Scores:
Pred. No.: 1.05e-32 Length: 639
Score: 395.50 Matches: 82
Percent Similarity: 85.8% Conservative: 9
Best Local Similarity: 77.4% Mismatches: 10
Query Match: 74.2% Indels: 5
Gaps: 2
DB: 9
US-10-715-129-2 (1-106) x CC963171 (1-639)

Qy 4 GlyArgAsnAraGlyArgAsnMetSerPheArgSerSerPro-----GluLys 19
Db 317 GGAAGAAACAACCGAAGAAACGTAAACTTTCGCGTCCACCGGCAATGACGTGAGAAC 258

Qy 20 SerLyGInGluLeuAraGArgSerPheSerAlaGlnLysArgMetCilieProAlaAsn 39

Pred. No.: 2.02e-25 Length: 471
 Score: 327.00 Matches: 72
 Percent Similarity: 82.2% Conservative: 11
 Best Local Similarity: 71.3% Mismatches: 12
 Query Match: 61.4% Indels: 6
 DB: 1 Gaps: 3

US-10-715-129-2 (1-106) x AVS44303 (1-471)

QY 1 MetAspVal---GlyAArgAsnAsnArgLysAsnMetSerPheArgSerSer----- 16
 DB 342 ATGACGCGAGAGGAGATCATACCGGAAACACGAGTTTTCGTTCAGCTCCAGCT 283
 QY 17 ProGluLysSerLysGlnGluLeuArgSerPheSerAlaGlnLys-----ArgMet 34
 DB 282 CGGATTATGGGAGGAGCAAAATGTTTGGACATTGTGTGCGAGAACAGTCCAGGAGG 223
 QY 35 MetIleProAlaMetLysPheSerLeuGlnSerLeuPheLeuValGlyLeuThrAla 54
 DB 222 CTATATACGAGAGTACTTCACTTACATCAATGATGCTTGTGTTGCTTCACAGCA 163
 QY 55 SerLeuLeuIleLeuProLeuValLeuProLeuProProProProPheMetLeuLeu 74
 DB 162 TCTCTCTGATCTTACCGTGTATCTTCCACCATGCTCCTCCTTTATGCTGCTT 103
 QY 75 LeuValProIleGlyIleMetValLeuValLeuAlaPheMetProSerSerHis 94
 DB 102 TTGATCTCATTTGGAGATTATGGTTTGGTTTGTCTTATGCTTCAATCCCTTATAT 43
 QY 95 Ser 95
 DB 42 TCC 40

RESULT 7
 BH528419/c 616 bp DNA linear GSS 13-DEC-2001
 LOCUS BH528419
 DEFINITION B0GCJ46TR B0GC Brassica oleracea genomic clone B0GCJ46, genomic survey sequence.

ACCESSION BH528419
 VERSION BH528419.1 GI:17736504
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE
 AUTHORS Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 616)

TITLE
 JOURNAL Aylee, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Uterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
 Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis Genome Res. 15 (4), 487-495 (2005)
 PUBMED 15805490

COMMENT
 OTHER GSSs: B0GCJ46TR
 CONTACT: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source location/Qualifiers

1..616
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone="B0GCJ46"
 /note="Vector: pHOSt. Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"

ORIGIN

Alignment Scores:
 Pred. No.: 1.8e-21 Length: 616
 Score: 292.00 Matches: 66
 Percent Similarity: 75.7% Conservative: 18
 Best Local Similarity: 59.5% Mismatches: 19
 Query Match: 54.8% Indels: 8
 DB: 9 Gaps: 3

US-10-715-129-2 (1-106) x BH528419 (1-616)

QY 1 MetAspValGlyAArgAsnAsnArgLysAsnMetSerPheArgSerSerProGluLysSer 20
 DB 534 ATGACGCGAGAGGAGATCATACCGGAAACACGAGTTTTCGTTCAGCTCCAGCT 484
 QY 21 LysGlnGluLeuAArgSerPheSerAlaGlnLys-----ArgMetMetIleProAla 38
 DB 483 AAGCAAGATATTTTCCGACATTGTCTCTAGAACAGTCCGAGAGGCTATTCAGCG 424
 QY 39 AsnTyrPheSerLeuGlnSerLeuPheLeuValGlyLeuThrAlaSerLeuLeu 58
 DB 423 AGTACTTCACTTGGAGTCAATGATGCTTCTTGTGCTTCACAGCATCGCTTGATC 364
 QY 59 LeuProLeuValLeuProProLeuProProProProPheMetLeuLeuValProIle 78
 DB 363 CTTCCTGATCTTCCACCGTGTCTTCTCCTCCGTCATGCTGCTTCAATTCATAT 304
 QY 79 GlyIleMetValLeuValLeuValAlaPheMetProSerSerHisSerAsnAlaAsn 98
 DB 303 GCGATATGCTTGTGCTTATGCTTCTTCTTATGCTTCTTCTTCTTCTTATTTCT 244
 QY 99 ThrAsp-----ValThrCysAsnMet 106
 DB 243 TCTATGCCAAACATGTAACAGAACTTATATG 211

RESULT 8
 BZ039389/c 793 bp DNA linear GSS 09-OCT-2002
 LOCUS BZ039389
 DEFINITION Ik20F06.g1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.

ACCESSION BZ039389
 VERSION BZ039389.1 GI:23626714
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE
 AUTHORS Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 793)

TITLE
 JOURNAL Delehanty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea unpublished (2002)
 CONTACT: Richard K. Wilson

COMMENT
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.wustl.edu
 Plate: Ik20 row: f column: 06
 Seq primer: -28Rpot reverse
 Class: shotgun
 High quality sequence start: 77
 High quality sequence stop: 511.
 Location/Qualifiers

FEATURES
 source location/Qualifiers

1..793
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone="B.oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000D3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic

DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Alignment Scores:

| Pred. No.: | 2,22e-19 | Length: | 793 |
|------------------------|----------|---------------|-----|
| Score: | 274.00 | Matches: | 64 |
| Percent Similarity: | 77.6% | Conservative: | 19 |
| Best Local Similarity: | 59.8% | Mismatches: | 17 |
| Query Match: | 51.4% | Indels: | 8 |
| DB: | 9 | Gaps: | 2 |

US-10-715-129-2 (1-106) x BZ039389 (1-793)

```

Qy 5 ArgAsnAsnArgLyAsnMetSerPheArgSerSerProGlu-----LysSerLys 21
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 610 AGAGATTAACGACAGCATGCCCCTTCGGGAGACCACTCAAGATCAGATGATGAGTACG 552
Qy 22 GlnGlnLeuArgArgSerPheSerAlaGln-----LysArgMetMetIleProAla 38
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 551 CAGGAATATTACAGACATGTCCTCTCAGAACAGTACGAGAGGCTAATTAATCG 492
Qy 39 AsnTrpSerLeuGlnSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuIle 58
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 491 AGTACTTCAGTTGGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
Qy 59 LeuProLeuValLeuProProLeuProProProProPheMetLeuLeuValProIle 78
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 431 CTTCGGTGAATCTTCCACCGTGCCCTCCCTCCCTCAAGCTGCTGCTGCTGCTGCTATT 372
Qy 79 GlyLeuMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsn 98
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 371 GGGATTATGGTTTGGCTTATGCTTCTGCTTAAAGCTTCTCTATATGCCAAACATGTA 312
Qy 99 ThrAsp-ValThrCysAsn 104
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 311 ACAGACGACCTTACATGTAAC 293

RESULT 9
LOCUS BGI25370 382 bp mRNA linear EST 31-JAN-2001
DEFINITION BGI25370 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CTOPGM10 5' sequence, mRNA sequence.
ACCESSION BGI25370
VERSION BGI25370.1 GI:12625558
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 382)
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Uterback, T.,
Hansen, C., Roming, C. and Tanksley, S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
source
1..382
location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="CTOPGM10"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/clone_id="tomato shoot/meristem"

```

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Small expanding leaves from the growing tip were taken from greenhouse plants (4-6wks old TA96). Tissue was immediately frozen in liquid nitrogen."

ORIGIN

Alignment Scores:

| Pred. No.: | 2,17e-19 | Length: | 382 |
|------------------------|----------|---------------|-----|
| Score: | 270.50 | Matches: | 56 |
| Percent Similarity: | 77.9% | Conservative: | 18 |
| Best Local Similarity: | 58.9% | Mismatches: | 16 |
| Query Match: | 50.8% | Indels: | 5 |
| DB: | 2 | Gaps: | 2 |

US-10-715-129-2 (1-106) x BGI25370 (1-382)

```

Qy 13 PheArgSerSerProGlnLysSerLyGlnGlnLeuArgArgSerPheSer---AlaGln 31
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 353 TTCACACACATTATGACAGGCATTAAGATGACATTAAGATCATTTACACACACAT 294
Qy 32 LysArgMetMetIleProAlaAsnTrpSerLeuGlnSerLeuPheLeuValGly 51
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 293 GGCAGAAAATGTTGCAATGAATTAATTTAGCTTGAAGTCAATTAATTTGTTACTGGT 234
Qy 52 LeuThrAlaSerLeuLeuLeuLeuProLeuValLeuProProLeuProProPhe 71
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 233 CTTCACAGCATCTTTGTTACTTTGTCATGATGATCTTCACCATTCGACACACACCTTT 174
Qy 72 MetLeuLeuLeuValProIleGlyIleMetValLeuLeuValValLeuAlaPheMetPro 91
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 173 ATGTTGTTTACTTCCCAATTTTCATTCCTGTTGTTCTTATGATCTTATGCTTTATGCC 114
Qy 92 SerSerHisSerAsnAlaAsnThrAspValThrCysAsnPheMet 106
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 113 TCTAATGTTAGGAAAT-----GTGACTTGCTCATATCTT 81

RESULT 10
LOCUS AI774560 624 bp mRNA linear EST 18-MAY-2001
DEFINITION BGI255660 tomato resistant, Cornell Lycopersicon esculentum cDNA
clone CLBR12N17, mRNA sequence.
ACCESSION AI774560
VERSION AI774560.1 GI:5272601
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 624)
D'Accenzo, M., He, X., Lyman, J., Materu, A.L., Vialon, T., Holt, I.E.,
Liang, F., Upton, J., Roming, C.M., Craven, M.B., Fujii, C.Y.,
Bowman, C.L., Nieman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
Giovannoni, J.J. and Martin, G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
Other ESTs: TC2160
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html

FEATURES
source
1..624
location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="RI1-12 (35S)::Pro in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLBR12N17"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/ab_host="SOLR"

```


ORIGIN

/clone_11b="tomato resistant, Cornell"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLBR - Tomato Pseudomonas resistant EST library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

Alignment Scores:

| Pred. No.: | 3 97e-19 | Length: | 624 |
|------------------------|----------|---------------|-----|
| Score: | 270.50 | Matches: | 56 |
| Percent Similarity: | 77.9% | Conservative: | 18 |
| Best Local Similarity: | 58.9% | Mismatches: | 16 |
| Query Match: | 50.8% | Indels: | 5 |
| DB: | 1 | Gaps: | 2 |

US-10-715-129-2 (1-106) x AT774560 (1-624)

Qy 13 PheargSerSerProGluLysSerLysGlnGluLeuArgArgSerPheSer---AlaGln 31
Db 332 TTCAACAACATTATGAGGAGCAATGATGAGCATAGATCATTTACACAGGACAT 391
Qy 32 LysArgMetMetIleProAlaAsnTyrPheSerLeuGlnSerLeuPheLeuValGly 51
Db 392 GGCAAGAAATGTTGTCATGAAATATTATTAGCTTAGAGTCATATTATTGTTACTGGT 451
Qy 52 LeuThrAlaSerLeuLeuIleLeuProLeuValLeuProProLeuProProProPhe 71
Db 452 CTTCACAGCATCTTGTTCATGCTTTCGATGATGCTTCACACATTCACACACACCTTTT 511
Qy 72 MetLeuLeuValProIleGlyIleMetValLeuLeuValIleValAlaPheMetPro 91
Db 512 ATGTTGTTGTTAGTCCCAATTTTCATTTCTTTGTTCTTATGATCTTAGCTTTATGCTT 571
Qy 92 SerSerHisSerAsnAlaAsnThrAspValThrCysAsnPhMet 106
Db 572 TCTAATGTTAGGAAAT-----GTGACTGCTCATATCTT 604

RESULT 11

LOCUS BM688619 678 bp mRNA linear EST 27-APR-2005
DEFINITION BM688619 Lycopersicon esculentum cv. Micro-Tom fruit Lycopersicon
esculentum cDNA clone FC12CF10 5', mRNA sequence.
ACCESSION BM688619
VERSION BM688619.1 GI:62931030
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
AUTHORS Tenguete, T., Watanabe, M., Yano, K., Suzuki, H., Sakurai, N. and
Shibata, D.
TITLE Expressed sequence tags of full-length cDNA clones prepared from
the laboratory-grown miniature tomato (Lycopersicon esculentum)
cultivar Micro-Tom
JOURNAL Unpublished (2005)
COMMENT Contact: Daisuke Shibata
Kazusa DNA Research Institute;
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3947
Fax: 81-438-52-3948
Email: shibata@kazusa.or.jp.
Location/Qualifiers

FEATURES

source

1..678
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Micro-Tom"
/db_xref="taxon:4081"
/clone="FC12CF10"
/tissue_type="fruit"
/clone_11b="Lycopersicon esculentum cv. Micro-Tom fruit"

ORIGIN

Alignment Scores:

| Pred. No.: | 6.91e-18 | Length: | 678 |
|------------------------|----------|---------------|-----|
| Score: | 259.50 | Matches: | 53 |
| Percent Similarity: | 76.8% | Conservative: | 17 |
| Best Local Similarity: | 55.8% | Mismatches: | 20 |
| Query Match: | 48.7% | Indels: | 5 |
| DB: | 5 | Gaps: | 2 |

US-10-715-129-2 (1-106) x BM688619 (1-678)

Qy 13 PheargSerSerProGluLysSerLysGlnGluLeuArgArgSerPheSer---AlaGln 31
Db 359 TTCAACAACATTATGAGGAGCAATGATGAGCATAGATCATTTACACAGGACAT 418
Qy 32 LysArgMetMetIleProAlaAsnTyrPheSerLeuGlnSerLeuPheLeuValGly 51
Db 419 GGCAAGAAATGTTGTCATGAAATATTATTAGCTTAGAGTCATATTATTGTTACTGGT 478
Qy 52 LeuThrAlaSerLeuLeuIleLeuProLeuValLeuProProLeuProProProPhe 71
Db 479 CTTCACAGCATCTTGTTCATGCTTTCGATGATGCTTCACACATTCACACACACCTTTT 538
Qy 72 MetLeuLeuValProIleGlyIleMetValLeuLeuValIleValAlaPheMetPro 91
Db 539 ATGTTGTTGTTAGTCCCAATTTTCATTTCTTTGTTCTTATGATCTTAGCTTTATGCTT 598
Qy 92 SerSerHisSerAsnAlaAsnThrAspValThrCysAsnPhMet 106
Db 599 TCTAATGTTAGGAAAT-----GTGACTGCTCATATCTT 631

RESULT 12

LOCUS B1934590/c 606 bp mRNA linear EST 18-OCT-2001
DEFINITION B1934590 tomato flower, anthesis Lycopersicon esculentum cDNA
clone CTOD20B6 5' end, mRNA sequence.
ACCESSION B1934590
VERSION B1934590.1 GI:16249062
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
AUTHORS van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,
Uteback, T., Van Aken, S., Ronning, C.M., Niemann, W., Fraser, C.M.,
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato flower tissue, anthesis (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers

FEATURES

source

1..606
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOD20B6"
/tissue_type="flower"
/dev_stage="anthesis"
/clone_11b="tomato flower, anthesis"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosids; eurosids II; Sapindales; Rutaceae; Citrus.

AUTHORS

1 (bases 1 to 657)
Close, T.J., Roose, M.L., Arpaia, M.L., Federici, C.F., Fenton, R.D.,
Wanamaker, S., Focht, E., Sievert, J., Robinson, P., Kim, H.R.,
Kudrna, D., Stum, D., Yost, D., and Wing, R.

TITLE

Development of EST Resources and New Genetic Markers for California
Citrus - Washington Naval Orange Stored Fruit Rind

JOURNAL

Unpublished (2004)

COMMENT

Contact: Timothy Close
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124
Tel: 9097873318
Fax: 9097874437
Email: timothy.close@ucr.edu
Seq primer: T3.

FEATURES

Location/Qualifiers

1..657
/organism="Citrus sinensis"
/mol_type="mRNA"
/cultivar="Washington navel"
/db_xref="taxon:2711"
/clone="CS_MBC0006M17"
/tissue_type="Rind"
/dev_stage="Commercially producing trees"
/lab_host="E. coli TUC121"
/clone_id="Washington Naval Orange Stored Fruit Rind cDNA
Library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Fruits were collected
January-March 2003 (Federici, Roose lab, Focht, Sievert &
Robinson, Arpaia lab). Four samples related to storage
conditions were produced: 1) Fresh-picked in Mentone
(Arnott Brothers Enterprises, Mentone, CA), 2) after 21
days storage at 5C at Kearney then transported to UC
Riverside on ice, 3) after 5 additional days storage at
11C at Kearney, sampled immediately, 4) fruit grown in
southern CA were obtained from Redlands foothill Packing
House after commercial packing, X-ray irradiated at 300 Gy
by Surebeam, then stored 1 day at ambient temperature.
Rind tissue (juice vesicles) were collected. Tissues were
snap frozen and then stored at -80C until further
processing. Fenton (Close lab) purified RNA by the phenol
method described in J. Japanese Soc. Hort. Sci. 1996. 64
(4): 809-814, purified poly(A) mRNA using a PolyATTrack
mRNA Isolation System IV (Promega), produced a primary
cDNA library using a lambda ZAP XR cDNA Synthesis Kit
(Stratagene), then mass-excised one million pfu from the
primary library to produce a phagemid population.
Phagemids were plated, plasmid DNA purified, cDNA clones
archived, and DNA sequences determined bi-directionally
using an ABI3730 at the Arizona Genomics Institute,
University of Arizona (Kim, Kudrna, Stum, Yost, Wing).
Chromatogram files were downloaded by FTP to UC Riverside
(by Close), then processed at UC Riverside (by Wanamaker,
Close lab), using the HarVEST pipeline
(http://harvest.ucr.edu) to remove vector and cloning
oligo sequences and various contaminants, and to trim to a
high quality region. Sequences that retained a phred 17
region of at least 100 bases were deposited to GenBank."

ORIGIN

Alignment Scores:

Pred. No.: 1..16-17 Length: 657
Score: 257.50 Matches: 56
Percent Similarity: 75.6% Conservative: 12
Best Local Similarity: 62.2% Mismatches: 17
Query Match: 48.3% Indels: 5
DB: 7 Gaps: 2

US-10-715-129-2 (1-106) x CN192376 (1-657)

Qy 18 GlutyserryGIngluLeuAArgSerPheSer---AlaGlnlyArgMetMetIle 36
Db 47 GAGAGAGAGAAAGTGTGATACCGAGATTAATCTCGAAGGCGCATCCAGAAACGTGTT 106
Qy 37 ProhlaanTyPheSerLeuGluSerLeuPheLeuValGlyLeuThraIAserLeu 56
Db 107 TCGGCAAGCATATTCACCGTAGATCAATGCTTTTGCTGATATGCTACGCGCCATG 166
Qy 57 LeuileuProLeuValLeuProProLeuProProLeuProProMetLeuLeuVal 76
Db 167 CTGATCCGTCATTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTT 226
Qy 77 ProileGlyIleMetValLeuLeuValValLeuIlePheMetProSerSerHisSerAsn 96
Db 227 CTTATGATTTCTTACCGCTTCCTTGTGCTTGGCATTCATTCCTCTTAAT----- 277
Qy 97 AlaAnThraPValThrCysAsnPheMet 106
Db 278 ---CTAAGAGTATTAATCTTCACGTACGTG 304

RESULT 15
DN625498 749 bp. mRNA linear EST 12-APR-2005
LOCUS UCRCA01.06113 x Bark of Madame Vinous Sweet Orange Scion on
DEFINITION Standard Sour Orange Rootstock, Citrus Tristeza Virus Challenged -
ACCESSION UCRCA01 Citrus aurantium cDNA clone CA_SBA006113, mRNA sequence.
VERSION DN625498.1 GI:61694598
KEYWORDS EST.
SOURCE Citrus aurantium
ORGANISM Citrus aurantium

REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
Rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 749)

AUTHORS

Close, T.J., Roose, M.L., Federici, C.F., Mandal, J., Fenton, R.D.,
Kudrna, D., Wallen, C., Patino, A., Wanamaker, S., Kim, H.R.,
Polek, M.L., Collins, K., Wisniewski, M., Byrne, M., Stum, D., Smart, D.,
Muller, C., and Wing, R.

TITLE

Development of EST Resources and New Genetic Markers for California
Citrus - Bark of Madame Vinous Sweet Orange Scion on Standard Sour
Orange Rootstock, Challenged with Citrus Tristeza Virus cDNA
Library UCRCA01

JOURNAL

Unpublished (2005)

COMMENT

Contact: Timothy J. Close
Department of Botany & Plant Sciences
University of California
Riverside, CA 92521-0124, USA
Tel: 909-787-3318
Fax: 909-787-4437
Email: timothy.close@ucr.edu
Seq primer: T3.

FEATURES

Location/Qualifiers

1..749
/organism="Citrus aurantium"
/mol_type="mRNA"
/cultivar="Standard Sour Orange"
/db_xref="taxon:43166"
/clone="CA_SBA006113"
/tissue_type="bark below bud union, predominantly C.
aurantium"
/dev_stage="young trees"
/lab_host="E. coli TUC121"
/clone_id="Bark of Madame Vinous Sweet Orange Scion on
Standard Sour Orange Rootstock, Citrus Tristeza Virus
Challenged - UCRCA01"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; All plants were grown and
inoculated at the Central California Tristeza Bradication
Agency (CCTBA), under the direction of Marylou Polek, with
assistance by Cindy Wallen, who grew and inoculated the
plants, and Tony Patino who collected tissue and ran the
ELISA tests. All plants were Madame Vinous sweet orange

(*Citrus sinensis* Osbeck) budded onto standard Sour Orange rootstock (*Citrus aurantium* L.). They were inoculated onto the scion with bark patches infected with various strains of CTV. Two inoculation dates: 1/24/03, 7/14/04. Three collection dates: 8/19/04 (2 sets collected on this date), 10/21/04, and 12/9/04. Four plant 'days after inoculation' groups: 36, 99, 160, 573 days after inoculation. Number of plants collected for each set: 36-day (18 plants), 99-day (30 plants), 160-day (38 plants), and 573-day (18 plants). These included three virus-free control plants for each group, and 3 or more plants for each virus strain. Five quick-decline virus strains of CTV, identified by CTRBA isolate numbers, were used for the two sets of plants collected on 8/19/04: 107, 142, 96142, 96180, and 96208. Nine quick-decline virus strains of CTV were used for 10/21/04 and 12/9/04 collections: 107, 142, 96142, 96180, 96208, 160, 177, 146, 151. All of the 36-day plants failed to show positive ELISA tests of the scion before the sampling date, and continued to be negative in the rootstock when tested one and two months after the sampling date. The PCR test of the pooled RNA extracted from these plants was negative, so we had no convincing evidence that the virus was active at the bud union in these samples. Most of the 573 days-after-inoculation plants, which were growing outdoors, had tested ELISA positive for CTV at some time, but some of them gave negative tests shortly before sampling. However, they were showing decline symptoms, so all were judged to be infected. RNA was extracted from a pool of all the samples. This pool gave a strong RT-PCR positive result. ELISA tests indicated positive results for scion tissue from many of the plants of the 99-day and 160-day samples, but many were consistently negative. Some of the plants with negative ELISA scores did show some apparent symptoms of stunting, or stem pitting. RT-PCR was done by Claire Federici on some of the samples collected for the 99-day and 160-day samples, RNA from scion bark directly above bud-union sample piece. Those that were PCR tested had elevated ELISA scores, not quite positive, or they had symptoms that suggested they might be showing signs of CTV even though the ELISA titer did not indicate presence of the virus. Of the 27 99-day plants inoculated with CTV, 14 had positive ELISA tests. Of the 13 ELISA negative inoculated plants, 9 were tested by RT-PCR and 2 of these gave a positive result; the others were all negative by both types of test. Of the 35 160-day plants inoculated with CTV, 6 had positive ELISA tests. Of the 29 ELISA negative inoculated plants, 14 were tested by RT-PCR and all were negative. Which plants were used for the pools? All inoculated plants collected from the 36-day set and all inoculated plants from the 573-day set were used in two separate pools, regardless of ELISA scores. For the 99-day and 160-day plants, only ones with positive ELISA and/or RT-PCR results were used, and they were divided into four pools, with either high or low ELISA titer for each date. Samples were collected by Claire Federici, Mary Lou Polek, and Cindy Wallen. For each sample, a razor blade was used and the material collected in RNAlater (Ambion). The bud union with rootstock bark included the bark from just above the bud-union (this was the bottom cut of the scion bark piece) to about 1 cm below the bud union. That means this tissue included both rootstock bark and some scion bark. For every tree there were two tubes of bud union/rootstock. One bud union/rootstock tube of each tree was to be pooled for the cDNA library production. The other was to be saved for gene expression studies. There was one departure from this plan: the pools for the 160-day plants were made from both tubes of rootstock/bud union bark from these plants since the tissue amount was small. Samples in RNAlater were transported on ice to UC Riverside and then stored at -20 (until extracted. CTRBA contact information: Dr. Marylou Polek, Associate Plant Pathologist, Program Manager,

cctea@lightspeed.net, Department of Food and Agriculture, Plant Health & Pest Prevention Services, Integrated Pest Control Branch, Central California Tristeza Eradication Agency, 22847 Road 140, Tulare, CA 93274. Jayati Mandal and Raymond Fenton (Close lab) purified RNA by a TRIzol method, purified poly(A) mRNA using PolyAT Tract RNA Isolation Kit (Promega), produced a primary cDNA library using a lambda ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised 1 million pfu from the primary library to produce a phagemid population. The library was made from a mixture of equal quantities of RNA from each of the four sample types. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the Arizona Genomics Institute (Kim, Kidrne, Collura, Wisotski, Byrne, Stum, Smart, Muller, Wing). Chromatogram files were downloaded by FTP by Close, then processed by Manamaker (Close lab) using the Harvest pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were assembled, then chimeras were removed following manual inspection of assemblies (Manamaker, Close, Roose). Sequences that survived all removal steps were submitted to Genbank. Clones from this library are archived at the Arizona Genomics Institute (<http://www.genome.arizona.edu/orders/>).

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1,29e-17 | Length: | 749 |
| Score: | 257.50 | Matches: | 56 |
| Percent Similarity: | 75.6% | Conservative: | 12 |
| Best Local Similarity: | 62.2% | Mismatches: | 17 |
| Query Match: | 48.3% | Indels: | 5 |
| DB: | 8 | Gaps: | 2 |

US-10-715-129-2 (1-106) x DN625498 (1-749)

| | | | |
|----|-----|---|-----|
| QY | 18 | GIUlySerlyGlnIuLeuAgaRgSerPhSer---AlaGlnlySaRgMeMeIle | 36 |
| DB | 135 | GAGAAAGAGAAAGTTGATGATCAACGAAAGTTACTCCGAAAGGGCATCCGAAACCTGTT | 194 |
| QY | 37 | ProAlaAsnTyRheSerLeuGlnSerLeuPhaLeuLeuValGlyLeuThrAlaSerLeu | 56 |
| DB | 195 | TGGGCAAGCAATTACCCCTGGAAATCATGCTTGTCTGATGCTTGACGGCTCATTG | 254 |
| QY | 57 | LeuIleLeuProLeuValLeuProProLeuProProProPhaMetLeuLeuVal | 76 |
| DB | 255 | CTGATCTGCAATGATGCTTCGCCCTTGCCGCCCATTCGCTGCTTGGTT | 314 |
| QY | 77 | ProIleGlyIleMetValLeuLeuValValLeuAlaPhaMetProSerRhaSerAsn | 96 |
| DB | 315 | CCATATGATGATTTACCCGCTTGTGCTTGGCAATTCATGCTTCAAT | 365 |
| QY | 97 | AlaAsnThrAspValThrCysAsnPhMet | 106 |
| DB | 366 | ---GTAAAGATATTAATTCACGATGAGT | 392 |

Search completed: January 30, 2006, 11:17:32
 Job time : 3796 secs

ORIGIN

/translation="MDVGRNNRKNMFRSSPERSKDELRSFSQKRMIPANYSLE
SLFLVGLTASLILPLVPLPPPPMLLVPIGIVLVLVAFMPSHNNATDVIT
CNFM"

Query Match 100.0%; Score 732, DB 15, Length 732;
Best Local Similarity 100.0%; Pired. No. 2,6e-176;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCTCCATCTCCCTACTAGTCTGTTTTCACACAGTTTTCATCCAAACCAAC 60
DB 1 TTGCTCCATCTCCCTACTAGTCTGTTTTCACACAGTTTTCATCCAAACCAAC 60
QY 61 CAATACCAAAAGCTTCTCAAACTCCTCACTCAAAAGCTTCTTTCATCATGATCG 120
DB 61 CAATACCAAAAGCTTCTCAAACTCCTCACTCAAAAGCTTCTTTCATCATGATCG 120
QY 121 TTGAGTTAACTCGGATTTGTTTTCGATCTCTGTTTTCGAATCGTGGCCATCTTAATT 180
DB 121 TTGAGTTAACTCGGATTTGTTTTCGATCTCTGTTTTCGAATCGTGGCCATCTTAATT 180
QY 181 TTGTCGAATCTTCACCAATGCTTGATGCAAGTGCATTTGTTACACAGTGCCTTAA 240
DB 181 TTGTCGAATCTTCACCAATGCTTGATGCAAGTGCATTTGTTACACAGTGCCTTAA 240
QY 241 AGATCAGATCTTTGAGCAAAATTTTGTCACTGATCTTCTTAAATCCAAACGACAGCA 300
DB 241 AGATCAGATCTTTGAGCAAAATTTTGTCACTGATCTTCTTAAATCCAAACGACAGCA 300
QY 301 AAACAACCTCTGTGATGATTCGAAATCTCAAACTTACAAAAGATATTATTAACATT 360
DB 301 AAACAACCTCTGTGATGATTCGAAATCTCAAACTTACAAAAGATATTATTAACATT 360
QY 361 CAAGACGTTATTGAGCAACCGAGTATGACCTCGGAGAAACAACCGGAAAAATG 420
DB 361 CAAGACGTTATTGAGCAACCGAGTATGACCTCGGAGAAACAACCGGAAAAATG 420
QY 421 AGCTTTCGAAGTTGCGCGAGAAAAAGCAAGCAAGTATTCGCGCGAGCTTCTCGCGCAG 480
DB 421 AGCTTTCGAAGTTGCGCGAGAAAAAGCAAGCAAGTATTCGCGCGAGCTTCTCGCGCAG 480
QY 481 AAAAGATGATGATCCCGCGAATTAATTCAGTTAAGTCTCTGTTCTTAATGTTGGT 540
DB 481 AAAAGATGATGATCCCGCGAATTAATTCAGTTAAGTCTCTGTTCTTAATGTTGGT 540
QY 541 CTAACGGATCTCTGTTAATCTCCGTTAGTTTTCGCGCGGATTCCTCGCGCTCGCTTT 600
DB 541 CTAACGGATCTCTGTTAATCTCCGTTAGTTTTCGCGCGGATTCCTCGCGCTCGCTTT 600
QY 601 ATGCTGATATGTTCCCATTTGGATTATGTTTACCTCGCTTCTTCCCTCAATGCT 660
DB 601 ATGCTGATATGTTCCCATTTGGATTATGTTTACCTCGCTTCTTCCCTCAATGCT 660
QY 661 TCTTCTCATTTAAATGCTAATACAGATGTAACCTTGAATTTGATGTAATCTGAATTTA 720
DB 661 TCTTCTCATTTAAATGCTAATACAGATGTAACCTTGAATTTGATGTAATCTGAATTTA 720
QY 721 TTAATGATGAT 732
DB 721 TTAATGATGAT 732

RESULT 2
ATP24G16 100906 bp DNA linear PLN 16-APR-2005
LOCUS Arabidopsis thaliana DNA chromosome 3, BAC clone F24G16.
DEFINITION A138647
ACCESSION A138647
VERSION A138647.1 GI:6899904
KEYWORDS
ORGANISM Arabidopsis thaliana (chale crees)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

REFERENCE

1 D'Angelo, M., Vezzi, A., Modesto, D., Pigazzi, M., Valle, G.,
Nemes, H.M., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.
Unpublished
2 (bases 1 to 100906)
EU Arabidopsis sequencing project.
Direct Submission
Submitted (18-FEB-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
d'interet Public, Centre National de Sequencage - GENOSCOPS, 2 rue
Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr

COMMENT

Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES

1..100906
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/ecotype="Columbia"
1..1731
/note="overlap to BAC T16124, please refer EMBL Acc. no
AL138659 for analysis and annotation"
2..1888
/gene="F24G16.10"
/complement (42..1888)
/gene="F24G16.10"
complement (42..1888)
/gene="F24G16.10"
/note="Contains protein kinases signatures and profile
AA339-362; Protein kinases signatures and profile
AA458-470"
/product="receptor lectin kinase 3"
/protein_id="CA575793.1"
/db_xref="GI:7019668"
/db_xref="GGA:Q9M200"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR000985"
/db_xref="InterPro:IPR001220"
/db_xref="InterPro:IPR008271"
/db_xref="UniProt/TrEMBL:Q9M200"
/translation="MSHKVQLQIVLVLLTFSSTNSNGNFMERAAAAGLNGCYLLT
NTTKHSYGOAFNNTPVPINKSRSFNNIIRGIVPEHKKQSGHGMAYFSPTRGPGAS
DOYLGIFENETNNKANNVYALIELDKDSEFGDIDNHGIVINGLTGVASAGAY
DDEIDGPKKLSISTVMRLSIVSHTDQIANVTLIPATISVPKSLSLANDLSAPY
FLAEITLGTASTSGALGYVKNQSYSESVIYPAMDLYITPLPPYPKSYRTTRI
LAVCLTLAVFTALVASGIGVFYVRHKVKEVLEMBLQNGPFRPSYKELFNATKDFK
EKQLDKGSGQVYKGMLEPSDAELIAVKTSHSRQSMBSFLPSTIGLRHNPVR
LIGYCHKENLVLYVDFPMNGSLDRCLTSTNTSEBRLTWORFKIVATALLHL
HOMWGVUIVHRDILKPANVLLDHGMNARLDLPGIAKLYDGFQDQTSRVAGTGYIAP
LARTGASTSTDYARGVLMETVYCGRPLIRERAAENAVLDMITIELWESGGLPPAA
BSIRQSRNGRIETLVKLGILCAHTEILIRPMASVLOITLNVSHLPNNLIDVPAE
RLNGI"
complement (2..1888)
/gene="F24G16.10"
/number=1
2685..4651
/gene="F24G16.20"
complement (301..2685..4127,4214..4651)
/gene="F24G16.20"
complement (join(2685..4127,4214..4651))
/gene="F24G16.20"
/note="strong similarity to receptor lectin kinase 3 -
Arabidopsis thaliana, EMBL:U93161
Contains protein kinases signatures and profile
AA309-332; Protein kinases signatures and profile
AA425-437"


```
/codon_start=1
/product="receptor lectin kinase-like protein"
/protein_id="CAB75794.1"
/db_xref="GI:7019669"
/db_xref="GOA:Q9M129"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR000985"
/db_xref="InterPro:IPR001220"
/db_xref="InterPro:IPR008271"
/db_xref="UniProt/TREMBL:Q9M129"
/translation="MPEELKVLHIVLVLVLTLLSSSTYNSNGMTLEGSAA DNSIGDIT
LITNKKHSCQTFNNESIPIKDSFSPHFLGIVPEHTQSGHMSFVISEPTALPPA
SSQYGLFNSTGKSSNHYAIELDIQOEQEGDIDNHNVAWMLSIYSPTDPOO
LNTLFPABIPVPRKPLSLNRDISPLEMYGYTASGSI GAFHYMLSVATPK
VENDTFRIVPTLPYPKSSDRTKILAVCLLAFAPAVASGICPVPTTRKIK
EVLERNIQGPHFAFKELNATKDREKOLKGGGQVFGTLPSSNDELAVKKT
SHSQRQMSSEFLAISTIGRLRPNLVLGLGYCHKENLYLVDPFTNGSLDKYLDN
ENORLITWEORFKIKDVASALHDEWVOI IHRDIKPNVLIDHEMNAIDFGL
AKLYDGLDPTSRVAGTFGYIAPELRTGATSTVVAAGLVMLVWCGRMIEER
APEREVLVDMILIELMSGKLPDAESIQRQENRGRIELILKGLCAHHTELIRPN
MSAVMQLNGVSQLPDLIDVVRANLKGMPETSEVLGLNLTVSGTMTLINSFLSH
GR"
/complement(2685..4127)
/gene="P24G16.20"
/number=1
/complement(4128..4213)
/gene="P24G16.20"
/number=1
/complement(4214..4651)
/gene="P24G16.20"
/number=2
5498..8175
/gene="P24G16.30"
/complement(join(5498..5626,5840..5920,6042..6101,
6253..6332,6432..6483,6611..6748,7011..7274,7405..7462,
7549..7658,7819..8175))
/gene="P24G16.30"
/complement(join(5498..5626,5840..5920,6042..6101,
6253..6332,6432..6483,6611..6748,7011..7274,7405..7462,
7549..7658,7819..8175))
/note="several differences to EMBL:X81973
Contains ATP/GTP-binding site motif A (P-loop)
A169-176/Cysteine synthase/cystathionine beta-synthase
P-phosphate attachment site A143-161"
/codon_start=1
/product="cysteine synthase"
/protein_id="CAB75795.1"
/db_xref="GI:7019670"
/db_xref="GOA:Q43725"
/db_xref="UniProt/Swiss-Prot:Q43725"
/translation="MVAMIMASRPNRAKASQILSTLGNRSCTSMATSSALL
NPLTSSSSSTLRFRCPSEISLSFSSASDFSLAMRQSRSPADGSEBDPSVCEAV
KRTPGPGNLINADVSQLIGKTPMVLNLSIAKGCANVIAKLBIHBCSCVKGIGS
MTDABEOKGPISPKSYLVBPTSGNTGIGLAFIASRGVAILITMPASMEERVLK
ARGBALVITPRAKGTCAVQAKBEILKKTTPATYLLQFPDNPANPKIHRETTPRIWD
TKGKVDLFAVGIGTGITGVGRFKKSNPTQVIGVEPESDILSGKGGPKHIQGI
GAGETPNLDOKIMDEVIAISSEBAISTAKLAKBGLMWGISGAAANAIRKAKRP
ENAGKILAVTVNSMANVAFQVVPSPFSERYLSTPLFQSIREEVKKQPEV"
/complement(5498..5626)
/gene="P24G16.30"
/number=1
/complement(5627..5839)
/gene="P24G16.30"
/number=1
/complement(5840..5920)
/gene="P24G16.30"
/number=2
/complement(5921..6041)
/gene="P24G16.30"
/number=2
/complement(6042..6101)
/gene="P24G16.30"
```

```
/number=3
/complement(6102..6252)
/gene="P24G16.30"
/number=3
/complement(6253..6332)
/gene="P24G16.30"
/number=4
/complement(6333..6431)
/gene="P24G16.30"
/number=4
/complement(6432..6483)
/gene="P24G16.30"
/number=5
/complement(6484..6610)
/gene="P24G16.30"
/number=5
/complement(6611..6748)
/gene="P24G16.30"
/number=6
/complement(6749..7010)
/gene="P24G16.30"
/number=6
/complement(7011..7274)
/gene="P24G16.30"
/number=7
/complement(7275..7404)
/gene="P24G16.30"
/number=7
/complement(7405..7462)
/gene="P24G16.30"
/number=8
/complement(7463..7548)
/gene="P24G16.30"
/number=8
/complement(7549..7658)
/gene="P24G16.30"
/number=9
/complement(7659..7818)
/gene="P24G16.30"
/number=9
/complement(7819..8175)
/gene="P24G16.30"
/number=10
12111..18504
/gene="P24G16.40"

Query Match 100.0%; Score 732; DB 15; Length 100906;
Best Local Similarity 100.0%; Pred. No. 1, 9e-176;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGTCCTCATATTTCCCTAAGTCTGTTTCAACAGTTTCTTGATCCAAAC 60
DB 62556 TTGTCCTCATATTTCCCTAAGTCTGTTTCAACAGTTTCTTGATCCAAAC 62615
QY 61 CAATACCAAAAGCTTCAAACTCTTCAACCTCAAGCTTCTTCTTAACTTGAAATCG 120
DB 62616 CAATACCAAAAGCTTCAAACTCTTCAACCTCAAGCTTCTTCTTAACTTGAAATCG 62675
QY 121 TTGAGTTAACTGGATTTGTTTCTCATCTCTGTTTCTGAAATCGTGAGCCATCTTAATT 180
DB 62676 TTGAGTTAACTGGATTTGTTTCTCATCTCTGTTTCTGAAATCGTGAGCCATCTTAATT 62735
QY 181 TGTCGAATTTCTTCAACCAATGCTTCAAGCTGCAATGTTAAACCAATGCCCTAA 240
DB 62736 TGTCGAATTTCTTCAACCAATGCTTCAAGCTGCAATGTTAAACCAATGCCCTAA 62795
QY 241 AGATCAATCTTGAAGAAAATTTTGCACTGATCTTCTTAAATCCAAACGACACGCA 300
DB 62796 AGATCAATCTTGAAGAAAATTTTGCACTGATCTTCTTAAATCCAAACGACACGCA 62855
QY 301 AAACAACCTCTGTAGATGATTCGAGAATCTCAAACTTACAAAAGATATTATAACATT 360
DB 62856 AAACAACCTCTGTAGATGATTCGAGAATCTCAAACTTACAAAAGATATTATAACATT 62915
```

QY 361 CAAGACAGTTATTCGAACAAACGAGTCATGACCTCCGAAAGAAACAACCGAAAAACATG 420
Db 62916 CAAGACAGTTATTCGAACAAACGAGTCATGACCTCCGAAAGAAACAACCGAAAAACATG 62975
QY 421 AGCTTTGAAAGTTGCGCGAGAAAAGCAAGACAGATTACGCGAGATTCTCGCGCAG 480
Db 62976 AGCTTTGAAAGTTGCGCGAGAAAAGCAAGACAGATTACGCGAGATTCTCGCGCAG 63035
QY 481 AAAAGATGATGATCCCGGGAATTAATTCAGTTAGTCTCGTTCAATGGTTGCT 540
Db 63036 AAAAGATGATGATCCCGGGAATTAATTCAGTTAGTCTCGTTCAATGGTTGCT 63095
QY 541 CTACGCGCATCTCTGTTAATACCTTCGTTAGTTGCGCGCATCTCCGCTCCGTT 600
Db 63096 CTACGCGCATCTCTGTTAATACCTTCGTTAGTTGCGCGCATCTCCGCTCCGTT 63155
QY 601 ATGCTGCTATTGTTGCCCATTTGGGATTAAGTTTACTCGTGGTTCCTTCATGCT 660
Db 63156 ATGCTGCTATTGTTGCCCATTTGGGATTAAGTTTACTCGTGGTTCCTTCATGCT 63215
QY 661 TCTTCATTTCTAATGCTAATACAGATGTAATCTGCAATTCATGTAATCTGAATTTA 720
Db 63216 TCTTCATTTCTAATGCTAATACAGATGTAATCTGCAATTCATGTAATCTGAATTTA 63275
QY 721 TTATATGATGAT 732
Db 63276 TTATATGATGAT 63287

RESULT 3
AK176640
LOCUS
DEFINITION Arabidopsis thaliana mRNA for putative protein, complete cds,
clone: RAFL25-17-J07.

ACCESSION AK176640.1 GI:51971476
VERSION AK176640.1
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
Narusaka, M., Shin, T., Nakagawa, M., Sakamoto, N., Oishi, K.,
Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
Unpublished

TITLE
JOURNAL
AUTHORS
REFERENCE
2 (bases 1 to 747)
Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A.,
Narusaka, M., Shin, T., Nakagawa, M., Sakamoto, N., Oishi, K.,
Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K.,
Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
Direct Submission

COMMENT
Submitted (06-SEP-2004) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail: meeki@gsc.riken.jp,
URL: http://rarge.gsc.riken.jp/, Tel: 81-45-503-9625,
Fax: 81-45-503-9586)
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda phage-1-B vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This
clone is in a modified phage-script vector.
Please visit our web site (http://rarge.gsc.riken.jp/) for further
details.

FEATURES
SOURCE
1. 747
Location/Qualifiers

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAFL25-17-J07"
/ecotype="Columbia"
/note="common name: thale cress"
1. 747
/gene="At3g59900"
314..706
/gene="At3g59900"
/codon_start=1
/product="putative protein"
/protein_id="BAD4403.1"
/db_xref="GI:51971477"
/translation="MIREISNLOKDIINIDSYNNRVNDVGRNNRKNMFSRSPKRS
KQELRFSFAQKRMKIPANFSDSELPFLVGLTASLLIPLVPLPPPLLVPI
GIMVLIVLAFPSHSHNANTDVCNFM"

Query Match 99.7%; Score 730; DB 15; Length 747;
Best Local Similarity 100.0%; Pred. No. 8,3e-176;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GTCCTCCATTTCCCTACTAGTACTGTTTTCACAGATTCTTGATCCAAACCAACA 62
Db 1 GTCCTCCATTTCCCTACTAGTACTGTTTTCACAGATTCTTGATCCAAACCAACA 60
QY 63 ATACACAAAGCTTCTGAACTCCTTCACTCAAGCTTCTCTTTACATCTGAATGTT 122
Db 61 ATACACAAAGCTTCTGAACTCCTTCACTCAAGCTTCTCTTTACATCTGAATGTT 120
QY 123 GAGTTAACTCGGATTTGTTCTGCATCTCGTTTCTGAAATCGGGCCATCTTATTTG 182
Db 121 GAGTTAACTCGGATTTGTTCTGCATCTCGTTTCTGAAATCGGGCCATCTTATTTG 180
QY 183 TCTCGAATCTTCACCAATGCTTCGATCAAGCTGATGTTTAAACAGTTGCTTAAG 242
Db 181 TCTCGAATCTTCACCAATGCTTCGATCAAGCTGATGTTTAAACAGTTGCTTAAG 240
QY 243 ATCAGATCTTTGACCAAAATTTTGTGCTGATCTTCTTAAATCCAAACGACACGAAA 302
Db 241 ATCAGATCTTTGACCAAAATTTTGTGCTGATCTTCTTAAATCCAAACGACACGAAA 300
QY 303 ACAACCTCTGTATATGATTCGAGAAATCTCAACAAAGATTTATTAACATTC 362
Db 301 ACAACCTCTGTATATGATTCGAGAAATCTCAACAAAGATTTATTAACATTC 360
QY 363 AGACAGTTATTCGAACAAACGAGTCATGACGTCGGAAGAAACAACCGAAAAACATGAG 422
Db 361 AGACAGTTATTCGAACAAACGAGTCATGACGTCGGAAGAAACAACCGAAAAACATGAG 420
QY 423 CTTTCGAAGTTGCGCGAGAAAAGCAAGCAAGATTACGCGGAGTTTCTCGCGCAGAA 482
Db 421 CTTTCGAAGTTGCGCGAGAAAAGCAAGCAAGATTACGCGGAGTTTCTCGCGCAGAA 480
QY 483 AAGATGATGATCCCGCGGAATTAATTCAGTTAGTCTCTGTTCTTAATGGTTGCT 542
Db 481 AAGATGATGATCCCGCGGAATTAATTCAGTTAGTCTCTGTTCTTAATGGTTGCT 540
QY 543 AACGAGATCTGTATATCTTCGTTAGTTTGGCGCGTACTCCGCTCCGTTTAT 602
Db 541 AACGAGATCTGTATATCTTCGTTAGTTTGGCGCGTACTCCGCTCCGTTTAT 600
QY 603 GCTGCTATTGTTCCCATTTGGGATTAAGTTTACTGTCGTGTTCTTGCTTCATGCTTC 662
Db 601 GCTGCTATTGTTCCCATTTGGGATTAAGTTTACTGTCGTGTTCTTGCTTCATGCTTC 660
QY 663 TTTCTATTCTTAAGCTTAATACAGATGTAATCTTCAATTTCAATGAATTTAT 722
Db 661 TTTCTATTCTTAAGCTTAATACAGATGTAATCTTCAATTTCAATGAATTTAT 720
QY 723 ATATGATGAT 732

Db 721 ATATGATCAT 730

|||||

RESULT 4
LOCUS BT011724
DEFINITION Arabidopsis thaliana At3g59900 gene, complete cds.
ACCESSION BT011724
VERSION BT011724.1 GI:44917524
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shim,P., Carninci,P., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Narusaka,M., Sakurai,T., Satou,M., Seki,M., Shinozaki,K. and Ecker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 393)
TITLE Cheuk,R., Chen,H., Kim,C.J., Shim,P., Carninci,P., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Narusaka,M., Sakurai,T., Satou,M., Seki,M., Shinozaki,K. and Ecker,J.R.
JOURNAL Direct Submission
COMMENT Submitted (04-MAR-2004) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PONT (ORF) clones using the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Shim,P., Bowser,L., Chan,M.M., Chang,C.M., Dale,J.M., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
SOURCE location/Qualifiers
1..393
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U5594"
/ecotype="Columbia"
1..393
/note="unknown protein"
/codon_start=1
/product="At3g59900"
/protein_id="AAS49087.1"
/db_xref="GI:44917525"
/translation="MIRISNLQDIINIDSYNNRVMDVGNRRKMKMSFRSSPEKS
KQELRSFSAQKRMWIPANYPSSLFLVLTSLTLPVLPPLPPPPMLLVPI
GIVTLVLTAMPSSSHNATDVICNMF"

ORIGIN
Query Match 53.7%; Score 393; DB 15; Length 393;
Beet Local Similarity 100.0%; Pred. No. 1,4e-89;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

316 ATGATTCGAGAAATCTCAAACTTACAAAAGATATTATTAACATTCAGACAGTTATTCG 375
|||||

Db 1 ATGATTCGAGAAATCTCAAACTTACAAAAGATATTATTAACATTCAGACAGTTATTCG 60

Qy 376 AACACCGAGTCATGACGTCGGAAAGAACACCGGAAAAACATGAGTTTCGAAGTTCCG 435
|||||

Db 61 AACACCGAGTCATGACGTCGGAAAGAACACCGGAAAAACATGAGTTTCGAAGTTCCG 120
|||||

Qy 436 CCGGAGAAAAGCAAGCAAGATTACGGCGGAGTTTCGCGCGCAAGAAAGATGATGATC 495
|||||

Db 121 CCGGAGAAAAGCAAGCAAGATTACGGCGGAGTTTCGCGCGCAAGAAAGATGATGATC 180
|||||

Qy 496 CCGGCAATTAATTCAGTTTACGTCCTGTTCCATTTGTTGGTCTAACGCAATCTCTG 555
|||||

Db 181 CCGGCAATTAATTCAGTTTACGTCCTGTTCCATTTGTTGGTCTAACGCAATCTCTG 240
|||||

Qy 556 TTAATACCTCCGTTAGTTTGGCCGCCGTTACCTCCGCTCCGTTTATGCTGATTTGTT 615
|||||

Db 241 TTAATACCTCCGTTAGTTTGGCCGCCGTTACCTCCGCTCCGTTTATGCTGATTTGTT 300
|||||

Qy 616 CCCATTGGGATTAATGTTTACTGCTGCTTCCGCTTCATGACCTCTTCATTTAAT 675
|||||

Db 301 CCCATTGGGATTAATGTTTACTGCTGCTTCCGCTTCATGACCTCTTCATTTAAT 360
|||||

Qy 676 GCTAATACAGATGTACTGCAATTTTCATGTAA 708
|||||

Db 361 GCTAATACAGATGTACTGCAATTTTCATGTAA 393
|||||

RESULT 5
AY080817
LOCUS
DEFINITION Arabidopsis thaliana unknown protein (At2g44080) mRNA, complete cds.
ACCESSION AY080817
VERSION AY080817.1 GI:19423989
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Yamada,K., Liu,S.J., Sakano,H., Pham,P.K., Banb,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Shinozaki,K., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Full length cDNA Clones
JOURNAL Unpublished
COMMENT 2 (bases 1 to 904)
Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyer,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Banb,J., Chan,M.M., Chang,C.H., Chang,B., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Toriumi,M., Yu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,

Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
Meyers, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shin, P.,
Southwick, A., Davis, R. W., Ecker, J. R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC)
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.

FEATURES

source

Location/Qualifiers

1..904
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAF106-84-F13 (R11991)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector
(Lambda ZAP) as a XhoI/SacI insert."
1..904
/gene="At2g44080"
1..443
/gene="At2g44080"
444..851
/gene="At2g44080"
/codon_start=1
/evidence="experimental"
/product="unknown protein"
/protein_id="A087296.1"
/db_xref="GI:19423930"
/translation="MIREFSLQNDIINIOEHYSLNNMVDGPHRKNKTPFGSAPA
PIKMGKILPRTLSQNSPRLISASVYVLGTLSLILPLILPLPLPPPM
LLILPIGLVILVILAFMPSNKRHSSTPM"
852..904
/gene="At2g44080"

3' UTR

/gene="At2g44080"

ORIGIN

Query Match 24.9%; Score 182.6; DB 15; Length 904;

Best Local Similarity 70.0%; Pred. No. 9.2e-36; Indels 18; Gaps 3;

Matches 296; Conservative 0; Mismatches 109;

270 CTGATCTTTAAATCCAAACGACACAGCAAAACAACCTCTGATGATTCAGAGAAT 329
398 CTGATTTTCCAGATCAAACTCGAAAAGAGAAAAGCTCTTAATGATTCGTGAGTT 457
330 CTCAAATTTACAAAAGATTTAATCAATTCGAAGACATTTGAAACGAGATCAT 389
458 CTCAGCTCTCAAAAACGACATCAATTAACATTTCAAGAACTTCTCAACAACAT 517
390 GGAAGTGGAGA---AACAACGGAAAAACATGAGCTTCGAGTT-----CGCC 437
518 GGAAGTGAAGAGATCATTAACGGAAGAAACAGATTTCTGATGATTCAGCTCC 577
438 GGAAGAAAGACAGACAGATTCAGCGAGTTCTCGCGCAGAA-----AAGATGAT 491
578 GATTATGGGGAAGAAATTTGTTCCGACATTCGTCCGAGAACAGTCCAGAGAGCT 637
492 GATCCCGGCAATTTATTTCAATTAGAGTCTCTGTTCCATTTGTTGCTTAAACGATC 551
638 AATATACGCAATCTCACTTCAATTAAGATCAATGTTGTGTTGTGTCTCAACAGATC 697
552 TCTGTTAATCTCGTTAGTTTGGCGCGTTAAGCTCCGCTCCGTTAATGCTGATTT 611
698 TCTCTGATCTTACCGTTGATTTCTTCAACATTCCTCCTCTCTTTAATGCTGATTT 757
612 GGTTCATTTGGATTAATGTTTACTGCTGCTTCTTGACCTTCAAGCTTCTTCAATTC 671
758 GATTCATTTGGATTAATGTTTGTGTTGCTTATGCTTCTTCAATTC 817
672 TAA 674
818 CAA 820

RESULT 6
AC004005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL
TITLE
COMMENT
FEATURES
source

AC004005 109741 bp DNA linear PLN 11-MAR-2002
Arabidopsis thaliana chromosome 2 clone F6E13 map C1C10F02,
C1C02E07, complete sequence.
AC004005.3 GI:20197067
HTG.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Bakayocae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 109741)
Rounley, S.D., Kaul, S., Lin, X., Ketchum, K.A., Crosby, M.L.,
Brandon, R.C., Sykes, S.M., Mason, T.M., Kervilave, A.R., Adams, M.D.,
Somerville, C.R. and Venter, J.C.
Unpublished
2 (bases 1 to 109741)
Lin, X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 109741)
Town, C.D. and Kaul, S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@igf.org
On Apr 18, 2002 this sequence version replaced gi:5598402.
Location/Qualifiers
1..109741
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C10F02, C1C02E07"
/clone="F6E13"
/ecotype="Columbia"
/complement(1..4149)
/note="overlap with BAC clone F18019 (AC002333:1..4149)."
/rpt_family="AT_rich"
3302..4750
/gene="At2g43880"
/note="synonym: F6E13.1"
join(<3302..3748,3846..4135,4219..4447,4532..>4750)
/gene="At2g43880"
join(3302..3748,3846..4135,4219..4447,4532..4750)
/gene="At2g43880"
/codon_start=1
/product="putative polygalacturonase"
/protein_id="AAC23426.1"
/translation="MANNISPCAIIFPSINIFPLIKSSHPMPFNVORGARDDGRA
DAKSLTATNSLACGRARARVYVPGTYIVKLVYMGPKNITITKNDITLYAPNY
WDIGNSGYWTLFAKVRISVYGGTIDARIGATVCRKSGHCHQAGASISFSCNNYL
LSGLSFGNSGNMHTVYHSSNVRILEVRIAPGSGPRTDIIHQSSGASITSGDTAT
RIDGCIARLQSGSRNIWIERVNCGPGHGISISLGDVYANBEQVNTVYS8VPTGTOMV
RITMYRPSRSGFNWIERVRLINNNENPVYIDQYCPNKGCPROSGVYKIGVFA
NIKGTSTPIAMKLDGSGSNHCTGLRLODKLTYMRSASAYCRNMGRAAGVWVPPN
CM"

repeat_region
complement(5593..5613)
/rpt_family="AT_rich"
repeat_region
complement(6730..6753)
/rpt_family="AT_rich"
7192..8658
/gene="At2g43890"
/note="synonym: F6E13.2"
join(<7192..7632,7743..8032,8122..8350,8440..>8658)
/gene="At2g43890"
join(7192..7632,7743..8032,8122..8350,8440..8658)
CDS

Db 68067 GACGTGAGAGAGATCATTAACCGGAAAAACAGATTTCGTGTTCACTCCAGCTCC 68126
 QY 438 GGAGAAAGACAGAGAGATTACCGGAGTTTCTCGCGCAGAA-----AAGATGAT 491
 Db 68127 GATTATGGGAGAGAGAAATTGTTTCGACATTTCGTGCGAGACAGTCCAGAGAGCT 68186
 QY 492 GATCCCGGAGATTTATTTCACTTTAGAGTCTCTGTTCTTAATGTTGGTCTTAACGGCATC 551
 Db 68187 AATATGACGAGATCTTCACTTTAGATCAATGTTGTGCTGTTGTCTCAAGATC 68246
 QY 552 TCTGTTAACTCTCGTTAGTTTTCGCCGCTTACCTCCGCTCGTTATGCTGATTT 611
 Db 68247 TCTCTGATCTTACCGCTTGAATCTTCCACATTCGCTCTCTCTTTATGCTGCTTT 68306
 QY 612 GGTTCATGAGATTAATGTTTACTCGTGTCTTGCCTTACGCTTCTCTCATTC 671
 Db 68307 GATTCTATTGGAGTTATGTTTCTTATGTTTCTTATGTTCTTCAATTC 68366
 QY 672 TAA 674
 Db 68367 CAA 68369

RESULT 7
 BT000930 439 bp mRNA linear PLN 16-OCT-2002
 LOCUS Arabidopsis thaliana clone C105228 unknown protein (At2g44080)
 DEFINITION mRNA, complete cds.
 ACCESSION BT000930
 VERSION BT000930.1 GI:24030324
 KEYWORDS F1.1 CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 439)
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, B.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
 Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 Annotation based on January 2002 version of the Arabidopsis genome submitted to Genbank.
 TITLE Arabidopsis Open Reading Frame (ORF) Clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 439)
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, B.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission
 COMMENT Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 Annotation based on January 2002 version of the Arabidopsis genome submitted to Genbank.
 FEATURES
 source 1..439
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="2"
 /clone="C105228"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51."
 1..439
 /gene="At2g44080"
 1..408
 /gene="At2g44080"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AA041330.1"
 /db_xref="GI:24030325"

ORIGIN
 Query Match 22.8%; Score 167; DB 15; Length 439;
 Best Local Similarity 71.4%; Pred. No. 9.4e-32;
 Matches 269; Conservative 0; Mismatches 90; Indels 18; Gaps 3;
 3 UTR
 /gene="At2g44080"
 /translacion="MIREFSLQNDIINIIOEHSYLNNDVRDHRKNTSFRGSA
 PIMQKQELFRITLSSONSPPRLISASYSLSMVLVLGTRLSLLILPLIPPLPPPM
 LILIPGIVLMLVLAIFMPSNSKHSSTFM"
 409..439
 /gene="At2g44080"

RESULT 8
 AX506572 336 bp DNA linear PAT 27-SEP-2002
 LOCUS Sequence 1267 from Patent WO0216655.
 DEFINITION AX506572
 ACCESSION AX506572
 VERSION AX506572.1 GI:23387809
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1
 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 Stresses-regulated genes of plants, transgenic plants containing same, and methods of use
 Patent: WO 0216655-A 1267 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
 TITLE Arabidopsis thaliana
 JOURNAL The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
 REFERENCE 1
 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 Stresses-regulated genes of plants, transgenic plants containing same, and methods of use
 Patent: WO 0216655-A 1267 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG (CH)
 FEATURES
 source 1..336
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Query Match 18.1%; Score 132.6; DB 6; Length 336;
 Best Local Similarity 71.3%; Pred. No. 6e-23;
 Matches 209; Conservative 0; Mismatches 69; Indels 15; Gaps 2;
 13 GGAGATCATTAACCGGAAAAACAGATTTCGTGTTCACTCCAGCTCCGATTAATGGAG 72

and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Aikawa, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hamaguchi, K., Hirakawa, T., Horii, F., Iida, J., Imamura, K., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuyuhara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

FEATURES
source
1. 1302
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033100F02"

ORIGIN

Query Match 10.3%; Score 75.6; DB 15; Length 1302;
Best Local Similarity 66.7%; Pred. No. 2.1e-08;
Matches 108; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Oy 498 GGGCAATTATTAGTTAGTCTCTGTTCTTATGTTGGTGTCTAGCGCATCTGTT 557
|||
Db 865 GGGCAATTATTCTCGGTGAGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 924
|||
Oy 558 AATACTTCGTTAGTTTGGCGCGGTACCTCGCCTCGTTTATGCTGTTGCTTC 617
|||
Db 925 GATCTCTCCCTCGTCTGCTGCGCGCGCTGCCCGCCCGCGCATGCTGATGCTGCG 984
|||
Oy 618 CATTGGATTATGTTTACTGCTGCTGCTTCTGCTTATGCTC 659
|||
Db 985 GGTGGCATGCTGTGCTGCTGCTGCTGCGCGCTGCGCTTATGCTC 1026
|||

RESULT 11
LOCUS OSJN00249 46422 bp DNA linear PLN 16-APR-2005
DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNBA0058G03,
complete sequence.
ACCESSION AL731606
VERSION AL731606.2 GI:32482924
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
1 Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J.,
Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y.,
Wang, Q., Zhang, L., Lu, Y., Mu, Y., Lu, Y., Zhang, L. S., Yu, Z., Fan, D.,
Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J.,
Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R.,
Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J.,
Zhang, X., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q.,
Zhang, Y., Zhang, W., Wang, L., Ding, X., Sheng, C., Sun, Y., Chen, S.,
N. L., Zhu, F., Chen, W., Lan, H., Lai, Y., Cheng, Z., Gu, M., Jiang, J.,
Li, J., Hong, G., Xue, Y., and Han, B.
Sequence and analysis of rice chromosome 4
Nature 420 (6913), 316-320 (2002)

TITLE
JOURNAL
PUBMED
REFERENCES
AUTHORS

2 Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q., Hu, X.,
Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Wang, Q. J., Zhang, L.,
Liu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y., Li, C.,
Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K., Zhou, B.,
Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P., Fu, G.,
Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F., Tu, Y. F.,

TITLE

Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X. Y.,
Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W., Wang, L. J.,
Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L., Zhu, F. H. and
Hong, G. F.

REMARK

COMMENT

Submitted (04-MAY-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNBA0058G03.
On Jul 9, 2003 this sequence version replaced gi:20451715.
Web site: <http://www.ncgr.ac.cn>
----- Summary Statistics

Assembly program: phrap

----- This is a complete sequence.

Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (<http://www.softberry.com/>),
genscan (<http://CCR-081.mit.edu/GSNSCAN.html>), GenesMarkES
(<http://genemark.biology.gatech.edu/Genemark/>), tRNAscan-SR (Sean
Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SR/>), searches of the
complete sequence against NCBI non-redundant protein database (nr)
(<http://ncbi.nlm.nih.gov/dblast/>) and the EST database at NCGR.

FEATURES
source

1. 46422
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="japonica"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="4"
/clone="OSJNBA0058G03"
/clone_11b="CGI-OSJNBA"
/complement(1144..2100)
/gene="OSJNBA0058G03.1"
/complement(1144..2100)
/gene="OSJNBA0058G03.1"
/codon_start=1
/protein_id="CAE02427.2"
/db_xref="GI:38344583"
/db_xref="UniProt/TREMBL:Q7F9D9"
/translation="MPTFTNIGVTVIVLISVILGLICLGSINLPOLLKRR
RNYQANDOLSGFNGPWLRTILIVLIMWGKYLPLFNGBRFLSDQMWY
KRYVTSNGRPAFGFLAFLPLSALNQBNQGANRKNQCTICAVPLCCPSILY
EACVVFTHPIASNDQTSKVAKYTSASVHDGACTPFLSSIFLGFIVFLY
VLPVGGQILSLVINKGLRRIRYMLIFATGILFRAMFGSVLPWPGELVHESLVPS
FLVLMRAAMGIVILVYFPVAETFEVRNQHIELQTSISAL"
/complement(6127..6681)
/gene="OSJNBA0058G03.2"
/complement(6127..6681)
/gene="OSJNBA0058G03.2"
/codon_start=1
/protein_id="CAD41044.1"
/db_xref="GI:21742878"
/db_xref="GOA:Q7XU91"
/db_xref="InterPro:IPR001471"
/db_xref="UniProt/TREMBL:Q7XU91"
/translation="MASAAAGGCGYGVRRKRWSEIRPQTKRIWGSFE
SAAVAADVVALRLRGDAQNFPGSDVRLPRPASSPGDRAAALAAADVRREP
ALVAVGAAAAGELAAVMAGLVEVEQGLGSDDEFVDSRLVAEMVAEMLLDPP
VAAVDSBEGPCMAHSGIMDAC"
17958..18485
/gene="OSJNBA0058G03.3"
17958..18485
/gene="OSJNBA0058G03.3"
/codon_start=1
/protein_id="CAE02428.1"
/db_xref="GI:32482926"
/db_xref="GOA:Q7XRM1"
/db_xref="InterPro:IPR007087"
/db_xref="UniProt/TREMBL:Q7XRM1"

| ORIGIN | Query Match | 10.3%; | Pred. No. 6; | DB 15; | Length 46422; |
|--------------|--|--|---------------------|-----------|-----------------|
| | Best Local Similarity | 66.7%; | Score: No. 1.7e-08; | | |
| | Matches 108; | Conservative 0; | Mismatches 54; | Indels 0; | Gaps 0; |
| Qy | 498 | GGCGAATTATTTTCAGTTTAGAGCTCTCTGTCTTCTTATGTTGGTGTCTTAACGGCATCTCTGT | 557 | | |
| Db | 32928 | GGGCAAGTACTTCTCGGTGGAATCGCTGCTCTGCTGGTGGTGGCGGAGCGGGTCTGTT | 32987 | | |
| Qy | 558 | AATACTTCGGTTAGTTTGGCCGCGCTTACCTTCGCGCTTATGCTGATTCGTTCC | 617 | | |
| Db | 32988 | GAATCCCTCCGCTGTCGTGCGCCGCGCTGCGCGCGCGCGCTCGAATGATTCGTGCTCC | 33047 | | |
| Qy | 618 | CATTGGATTTATGTTTTCATCTGCTGTTCTTGGCTTCAATGCC | 659 | | |
| Db | 33048 | GGTGGGATGCTGTGTCGTGCTGCTGCGCGCTTCAATGCC | 33089 | | |
| RESULT 12 | | | | | |
| AC157500 | | | | | |
| LOCUS | AC157500 | 90825 bp | DNA | linear | PLN 18-FEB-2005 |
| DEFINITION | Oryza sativa (japonica cultivar-group) BAC clone OSUNB0056004, | | | | |
| ACCESSION | from chromosome 4, complete sequence. | | | | |
| VERSION | AC157500 | | | | |
| KEYWORDS | AC157500.1 GI:59933327 | | | | |
| SOURCE | HTG. | | | | |
| ORGANISM | Oryza sativa (japonica cultivar-group) | | | | |
| | Oryza sativa (japonica cultivar-group) | | | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | |
| | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; | | | | |
| | Ehrhartoideae; Oryzaceae; Oryza. | | | | |
| REFERENCE | 1 (bases 1 to 90825) | | | | |
| AUTHORS | McCombie,W.R., Zuhavern,T., Nascimento,L., Balija,V., Kramer,M., | | | | |
| | Spielgel,L., Miller,B., Muller,S., Katzenberger,F., Andrade,M.V., | | | | |
| | Palmer,L. and Bell,M. | | | | |
| | Oryza sativa (japonica cultivar-group) BAC clone OSUNB0056004, | | | | |
| | from chromosome 4, complete sequence | | | | |
| | Unpublished | | | | |
| | 2 (bases 1 to 90825) | | | | |
| | McCombie,W.R. | | | | |
| JOURNAL | Direct Submission | | | | |
| REFERENCE | Submitted (18-FEB-2005) Lita Annenberg Hazen Genome Center, Cold | | | | |
| AUTHORS | Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, | | | | |
| JOURNAL | NF 11724, USA | | | | |
| | This sequence was finished as follows unless otherwise noted: all | | | | |
| | regions were either double-stranded or sequenced with an alternate | | | | |
| | chemistry or covered by high quality data (i.e., phred quality >= | | | | |
| | 30); an attempt was made to resolve all sequencing problems, such | | | | |
| | as compressions and repeats; all regions were covered by at least | | | | |
| | one plasmid subclone or more than one M13 subclone; and the | | | | |
| | assembly was confirmed by restriction digest. Clone OSUNB0056004 | | | | |
| | overlaps clone OSUNB0027P08 (AL731593) from base 54646 to base | | | | |
| | 90825. The overlap is from base 1 to base 36180 on OSUNB0027P08. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..90825 | | | | |
| | /mol_type="Oryza sativa (japonica cultivar-group) " | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /cultivar="Nipponbare" | | | | |
| | /db_xref="taxon:39947" | | | | |
| | /chromosome="4" | | | | |
| | /clone="OSUNB0056004" | | | | |
| | /clone_11b="EcoRI" | | | | |
| | 36680..36760 | | | | |
| misc_feature | /note="The assembly is covered by high quality sequences | | | | |
| | derived from a transposed plasmid subclone." | | | | |
| misc_feature | 37336..37516 | | | | |
| | /note="The assembly is covered by high quality sequences | | | | |
| | derived from a transposed plasmid subclone." | | | | |
| misc_feature | 41013..41445 | | | | |

ORIGIN /note="The assembly is covered by high quality sequences derived from a transposed plasmid subclone."

Query Match 10.3%; Score 75.6; DB 15; Length 90825;
Best Local Similarity 66.7%; Pred. No. 1.6e-08;
Matches 108; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 498 GCGCAATTATTTAGTTAGTCTGTCTCTATTTAGTGTGCTACAGGATCTGT 557
Db 58953 GGGCACTACTTCTCGGTGAGTGTCTCTCTCTCTGTGCTGACGGCTGCTGT 59012
Qy 558 AATACTTCGTTAGTTTGTCCGCGTTACTCTCGCTCGTTATGCTGATTGTTCC 617
Db 59013 GATCTCTCCCTGCTGTGCTGCGCGCTGCCCGCGCCGCTCATGTGATGTGTTGCC 59072
Qy 618 CATTTGGATATGTTTACTCTGCTCTTCTTCTGCTTCAATGCC 659
Db 59073 GGTGGCATGCTGTGCTGCTGCTGCGCGCTGCGTTATGCC 59114

RESULT 13
AP008210_221
WPCOMMENT

Sequence split into 355 fragments LOCUS AP008210 Accession AP008210

| Fragment Name | Begin | End |
|---------------|---------|---------|
| AP008210_000 | 1 | 110000 |
| AP008210_001 | 100001 | 210000 |
| AP008210_002 | 200001 | 310000 |
| AP008210_003 | 300001 | 410000 |
| AP008210_004 | 400001 | 510000 |
| AP008210_005 | 500001 | 610000 |
| AP008210_006 | 600001 | 710000 |
| AP008210_007 | 700001 | 810000 |
| AP008210_008 | 800001 | 910000 |
| AP008210_009 | 900001 | 1010000 |
| AP008210_010 | 1000001 | 1110000 |
| AP008210_011 | 1100001 | 1210000 |
| AP008210_012 | 1200001 | 1310000 |
| AP008210_013 | 1300001 | 1410000 |
| AP008210_014 | 1400001 | 1510000 |
| AP008210_015 | 1500001 | 1610000 |
| AP008210_016 | 1600001 | 1710000 |
| AP008210_017 | 1700001 | 1810000 |
| AP008210_018 | 1800001 | 1910000 |
| AP008210_019 | 1900001 | 2010000 |
| AP008210_020 | 2000001 | 2110000 |
| AP008210_021 | 2100001 | 2210000 |
| AP008210_022 | 2200001 | 2310000 |
| AP008210_023 | 2300001 | 2410000 |
| AP008210_024 | 2400001 | 2510000 |
| AP008210_025 | 2500001 | 2610000 |
| AP008210_026 | 2600001 | 2710000 |
| AP008210_027 | 2700001 | 2810000 |
| AP008210_028 | 2800001 | 2910000 |
| AP008210_029 | 2900001 | 3010000 |
| AP008210_030 | 3000001 | 3110000 |
| AP008210_031 | 3100001 | 3210000 |
| AP008210_032 | 3200001 | 3310000 |
| AP008210_033 | 3300001 | 3410000 |
| AP008210_034 | 3400001 | 3510000 |
| AP008210_035 | 3500001 | 3610000 |
| AP008210_036 | 3600001 | 3710000 |
| AP008210_037 | 3700001 | 3810000 |
| AP008210_038 | 3800001 | 3910000 |
| AP008210_039 | 3900001 | 4010000 |
| AP008210_040 | 4000001 | 4110000 |
| AP008210_041 | 4100001 | 4210000 |
| AP008210_042 | 4200001 | 4310000 |
| AP008210_043 | 4300001 | 4410000 |
| AP008210_044 | 4400001 | 4510000 |
| AP008210_045 | 4500001 | 4610000 |
| AP008210_046 | 4600001 | 4710000 |

| | | |
|--------------|----------|----------|
| AP008210_047 | 4700001 | 4810000 |
| AP008210_048 | 4800001 | 4910000 |
| AP008210_049 | 4900001 | 5010000 |
| AP008210_050 | 5000001 | 5110000 |
| AP008210_051 | 5100001 | 5210000 |
| AP008210_052 | 5200001 | 5310000 |
| AP008210_053 | 5300001 | 5410000 |
| AP008210_054 | 5400001 | 5510000 |
| AP008210_055 | 5500001 | 5610000 |
| AP008210_056 | 5600001 | 5710000 |
| AP008210_057 | 5700001 | 5810000 |
| AP008210_058 | 5800001 | 5910000 |
| AP008210_059 | 5900001 | 6010000 |
| AP008210_060 | 6000001 | 6110000 |
| AP008210_061 | 6100001 | 6210000 |
| AP008210_062 | 6200001 | 6310000 |
| AP008210_063 | 6300001 | 6410000 |
| AP008210_064 | 6400001 | 6510000 |
| AP008210_065 | 6500001 | 6610000 |
| AP008210_066 | 6600001 | 6710000 |
| AP008210_067 | 6700001 | 6810000 |
| AP008210_068 | 6800001 | 6910000 |
| AP008210_069 | 6900001 | 7010000 |
| AP008210_070 | 7000001 | 7110000 |
| AP008210_071 | 7100001 | 7210000 |
| AP008210_072 | 7200001 | 7310000 |
| AP008210_073 | 7300001 | 7410000 |
| AP008210_074 | 7400001 | 7510000 |
| AP008210_075 | 7500001 | 7610000 |
| AP008210_076 | 7600001 | 7710000 |
| AP008210_077 | 7700001 | 7810000 |
| AP008210_078 | 7800001 | 7910000 |
| AP008210_079 | 7900001 | 8010000 |
| AP008210_080 | 8000001 | 8110000 |
| AP008210_081 | 8100001 | 8210000 |
| AP008210_082 | 8200001 | 8310000 |
| AP008210_083 | 8300001 | 8410000 |
| AP008210_084 | 8400001 | 8510000 |
| AP008210_085 | 8500001 | 8610000 |
| AP008210_086 | 8600001 | 8710000 |
| AP008210_087 | 8700001 | 8810000 |
| AP008210_088 | 8800001 | 8910000 |
| AP008210_089 | 8900001 | 9010000 |
| AP008210_090 | 9000001 | 9110000 |
| AP008210_091 | 9100001 | 9210000 |
| AP008210_092 | 9200001 | 9310000 |
| AP008210_093 | 9300001 | 9410000 |
| AP008210_094 | 9400001 | 9510000 |
| AP008210_095 | 9500001 | 9610000 |
| AP008210_096 | 9600001 | 9710000 |
| AP008210_097 | 9700001 | 9810000 |
| AP008210_098 | 9800001 | 9910000 |
| AP008210_099 | 9900001 | 10010000 |
| AP008210_100 | 10000001 | 10110000 |
| AP008210_101 | 10100001 | 10210000 |
| AP008210_102 | 10200001 | 10310000 |
| AP008210_103 | 10300001 | 10410000 |
| AP008210_104 | 10400001 | 10510000 |
| AP008210_105 | 10500001 | 10610000 |
| AP008210_106 | 10600001 | 10710000 |
| AP008210_107 | 10700001 | 10810000 |
| AP008210_108 | 10800001 | 10910000 |
| AP008210_109 | 10900001 | 11010000 |
| AP008210_110 | 11000001 | 11110000 |
| AP008210_111 | 11100001 | 11210000 |
| AP008210_112 | 11200001 | 11310000 |
| AP008210_113 | 11300001 | 11410000 |
| AP008210_114 | 11400001 | 11510000 |
| AP008210_115 | 11500001 | 11610000 |
| AP008210_116 | 11600001 | 11710000 |
| AP008210_117 | 11700001 | 11810000 |
| AP008210_118 | 11800001 | 11910000 |
| AP008210_119 | 11900001 | 12010000 |

39614. .39912,41472. .41811,41908. .41971,42052. .42273,
/42726. .42904)
/gene="OSJNB80027P08.8"
/codon_start=1
/protein_id="CAD40970.2"

Query Match 10.3%; Score 75.6; DB 15; Length 118959;
Best Local Similarity 66.7%; Pred. No. 1.6e-08;
Matches 108; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 498 GGGCAATTATTTCAGTTAAGTCTCTGTTCTATTGGTTGGTCTAAAGGATCTCTGTT 557
|||
Db 4308 GGGCAAGTACTTCTCGGTGAGTGTCTGCTCTCTGCTGTGAGCGGCTGCTGCT 4367
QY 558 AATACCTCCGTAGTTTGGCGCCGTTACTCCGCTCCGTTTATGCTGCTATTGTTCC 617
|||
Db 4368 GATCTCTCCCGCTGCTGCTGCGCGCGCTGCCCGCCCGCGCTGATGCTGCTGCC 4427
QY 618 CATTGGATTATGTTTACTCGTCTGCTTTCCTTCATGCC 659
|||
Db 4428 GTTGGCATGCTGTGCTGCTGCTGCGCGCTGCGCTCATGCC 4469

RESULT 15

LOCUS DQ099057 629 bp mRNA linear PLN 18-JUL-2005
DEFINITION Arachis stenosperma clone AS1RN2P12 microsatellite sequence.
ACCESSION DQ099057
VERSION DQ099057.1 GI:70779734
KEYWORDS

SOURCE
ORGANISM Arachis stenosperma

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

1 (bases 1 to 629)

Moretzsohn,M.C., Lecl.,L., Proite,K., Guimaraes,P.M.,
Leal-Bertoli,I.,S.C.M., Gimenes,M.A., Martins,W.S., Valle,J.F.M.,
Gratcapaglia,D. and Bertoli,I.D.J.

A microsatellite based, gene-rich linkage map for the AA genome of
Arachis (Fabaceae)
Theor. Appl. Genet. (2005) In press
2 (bases 1 to 629)

Bertoli,I.D.J.

Direct Submission

Submitted (17-JUN-2005) Laboratorio de Ineracoes Planta-Prague II,
Embrapa Recursos Geneticos e Biotecnologia, Parque Estacao
Biologica, Brasilia, DF 70770-900, Brazil

Location/Qualifiers

1. .629

/organism="Arachis stenosperma"

/mol_type="mRNA"

/db_xref="taxon:217475"

/clone="AS1RN2P12"

1. .629

/note="microsatellite"

/rpl_type=tandem

ORIGIN

Query Match 9.6%; Score 70.2; DB 15; Length 629;
Best Local Similarity 63.2%; Pred. No. 5.4e-07;
Matches 108; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 498 GGGCAATTATTTCAGTTAAGTCTCTGTTCTATTGGTTGGTCTAAAGGATCTCTGTT 557
|||
Db 209 GTTCGCTTACAGCTTAGGTACATTCATCTGTGTCTGTGAGAGTGTGATGCT 268
QY 558 AATACCTCCGTAGTTTGGCGCCGTTACTCCGCTCCGTTTATGCTGCTATTGTTCC 617
|||
Db 269 GTTCTTCCGCTGATGCTGCGCGCTACACCGCGCGCTGCTCTCTCTTCTCC 328
QY 618 CATTGGATTATGTTTACTCGTCTGCTTTCCTTCATGCCCTTCTCTCA 668

Db 329 GTTCGATCATGCGCGCCCTCATGTTTCTGCGCTTCTGCGCTTCCATCA 379

Search completed: January 30, 2006, 05:32:20
Job time : 3813 sec

THIS PAGE BLANK (U8PT0)

gene, ARGOS, from Arabidopsis, comprising fully defined 732 base pair sequence given in the specification, that encodes the amino acid sequence comprising fully defined 106 amino acids also given in the specification, or comprising their antisense nucleotide sequences. The nucleic acid is useful for regulating organ development including size control in transformed plants and plant cells. This sequence corresponds to the cDNA for the novel auxin-inducible gene ARGOS.

SQ Sequence 732 BP; 199 A; 170 C; 125 G; 238 T; 0 U; 0 Other;

| Query Match | Score | DB | Length |
|-------------|-------|----|--------|
| 100.0% | 732 | 14 | 732 |

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | TTGTCCTCATTTCCCTACTAGTACTGTGTCACACAGTTTCTTGATCCAAACCAAAAC | 60 |
| Db | 1 | TTGTCCTCCTCATTTCCCTACTAGTACTGTGTCACACAGTTTCTTGATCCAAACCAAAAC | 60 |
| Qy | 61 | CAATTCACAAAGCTTCTCAAACCTCCTTCACTCAAGCTTCTTCCCTTACATCTGAATCG | 120 |
| Db | 61 | CAATTCACAAAGCTTCTCAAACCTCCTTCACTCAAGCTTCTTCCCTTACATCTGAATCG | 120 |
| Qy | 121 | TTGAGTTAACTGGGAATTTGTCGCAATCCCTCTGTTCTGAATGCTGGGCAATCCTAATT | 180 |
| Db | 121 | TTGAGTTAACTGGGAATTTGTCGCAATCCCTCTGTTCTGAATGCTGGGCAATCCTAATT | 180 |
| Qy | 181 | TGTCCTCGAATTTCTTCCACCAATTCCTCGATTCAGAGCTGATTTGGTTAAACAGTTGCCCTAA | 240 |
| Db | 181 | TGTCCTCGAATTTCTTCCACCAATTCCTCGATTCAGAGCTGATTTGGTTAAACAGTTGCCCTAA | 240 |
| Qy | 241 | AGATCGAGTCTTTGAGCAAAATTTTGTCACTGATCTTCTPAATCCAAACCAAGCACAGCA | 300 |
| Db | 241 | AGATCGAGTCTTTGAGCAAAATTTTGTCACTGATCTTCTPAATTCCAAACCAAGCACAGCA | 300 |
| Qy | 301 | AAACAACCTCTGTAGATGATTCGAGAAATCTCAAACTTCAAAAAGATTTATPAACATT | 360 |
| Db | 301 | AAACAACCTCTGTAGATGATTCGAGAAATCTCAAACTTCAAAAAGATTTATPAACATT | 360 |
| Qy | 361 | CAAGACAGTTATTCGAACCAACCGAGTACATGACGTGGAAGAAACAAACGGAAAAACATG | 420 |
| Db | 361 | CAAGACAGTTATTCGAACCAACCGAGTACATGACGTGGAAGAAACAAACGGAAAAACATG | 420 |
| Qy | 421 | AGCTTCGAAGTTCCGCGGAGAAAAAGCAAGCAAGATTACGGCGGAGTTTCTCGCGCAG | 480 |
| Db | 421 | AGCTTCGAAGTTCCGCGGAGAAAAAGCAAGCAAGATTACGGCGGAGTTTCTCGCGCAG | 480 |
| Qy | 481 | AAAAAGATGATGATCCCGCGCAATTATTCAGTTTGAAGTCTCTGTTCCATTTGTTGGT | 540 |
| Db | 481 | AAAAAGATGATGATCCCGCGCAATTATTCAGTTTGAAGTCTCTGTTCCATTTGTTGGT | 540 |
| Qy | 541 | CTAACGGCATCTCTGTTAATCTTCGCGTAGTTTTCGCGCCGTTACCTCCGCTCCGTTT | 600 |
| Db | 541 | CTAACGGCATCTCTGTTAATCTTCGCGTAGTTTTCGCGCCGTTACCTCCGCTCCGTTT | 600 |
| Qy | 601 | ATGCTGCTATTGGTCCCATGGGAAATATGGTTTTACTGCTGCTTCCCTCAAGCCT | 660 |
| Db | 601 | ATGCTGCTATTGGTCCCATGGGAAATATGGTTTTACTGCTGCTTCCCTCAAGCCT | 660 |
| Qy | 661 | TCTTCTCATTTCTAATGCTAATAACAGATGTAATCTGCAATTCGATTAATCTGAATTTTA | 720 |
| Db | 661 | TCTTCTCATTTCTAATGCTAATAACAGATGTAATCTGCAATTCGATTAATCTGAATTTTA | 720 |
| Qy | 721 | TTATATGATGAT 732 | |
| Db | 721 | TTATATGATGAT 732 | |

| | |
|----------|----------------------------------|
| RESULT 2 | |
| AAC53386 | |
| ID | AAC53386 standard; DNA; 1005 BP. |
| XX | |
| AC | AAC53386; |
| XX | |

| | | | |
|----|--|--------------------------------|--|
| DT | 18-OCT-2000 | (first entry) | |
| XX | | | |
| DE | Arabidopsis thaliana | DNA fragment SEQ ID NO: 74346. | |
| XX | | | |
| KM | Hybridisation assay; genetic mapping; gene expression control; | | |
| KM | protein identification; signal transduction pathway; metabolic pathway | | |
| KM | promoter; termination sequence; ss. | | |
| XX | | | |
| OS | Arabidopsis thaliana. | | |
| XX | | | |
| PM | EP1033405-A2. | | |
| PD | 06-SEP-2000. | | |
| XX | | | |
| PF | 25-FEB-2000; 2000EP-00301439. | | |
| XX | | | |
| PR | 25-FEB-1999; | 99US-0121825P. | |
| PR | 05-MAR-1999; | 99US-0123180P. | |
| PR | 09-MAR-1999; | 99US-0123548P. | |
| PR | 23-MAR-1999; | 99US-0125788P. | |
| PR | 25-MAR-1999; | 99US-0126264P. | |
| PR | 29-MAR-1999; | 99US-0126785P. | |
| PR | 01-APR-1999; | 99US-0127462P. | |
| PR | 06-APR-1999; | 99US-0128234P. | |
| PR | 08-APR-1999; | 99US-0128714P. | |
| PR | 16-APR-1999; | 99US-0129845P. | |
| PR | 19-APR-1999; | 99US-0130077P. | |
| PR | 21-APR-1999; | 99US-0130449P. | |
| PR | 23-APR-1999; | 99US-0130510P. | |
| PR | 28-APR-1999; | 99US-0130891P. | |
| PR | 30-APR-1999; | 99US-0131449P. | |
| PR | 30-APR-1999; | 99US-0132048P. | |
| PR | 04-MAY-1999; | 99US-0132407P. | |
| PR | 05-MAY-1999; | 99US-0132484P. | |
| PR | 06-MAY-1999; | 99US-0132485P. | |
| PR | 06-MAY-1999; | 99US-0132486P. | |
| PR | 07-MAY-1999; | 99US-0132487P. | |
| PR | 11-MAY-1999; | 99US-0132863P. | |
| PR | 14-MAY-1999; | 99US-0134256P. | |
| PR | 14-MAY-1999; | 99US-0134218P. | |
| PR | 14-MAY-1999; | 99US-0134221P. | |
| PR | 14-MAY-1999; | 99US-0134370P. | |
| PR | 18-MAY-1999; | 99US-0134768P. | |
| PR | 19-MAY-1999; | 99US-0134941P. | |
| PR | 20-MAY-1999; | 99US-0135124P. | |
| PR | 21-MAY-1999; | 99US-0135533P. | |
| PR | 24-MAY-1999; | 99US-0135629P. | |
| PR | 25-MAY-1999; | 99US-0136021P. | |
| PR | 27-MAY-1999; | 99US-0136392P. | |
| PR | 28-MAY-1999; | 99US-0136782P. | |
| PR | 01-JUN-1999; | 99US-0137222P. | |
| PR | 03-JUN-1999; | 99US-0137528P. | |
| PR | 04-JUN-1999; | 99US-0137502P. | |
| PR | 07-JUN-1999; | 99US-0137724P. | |
| PR | 08-JUN-1999; | 99US-0138094P. | |
| PR | 10-JUN-1999; | 99US-0138540P. | |
| PR | 10-JUN-1999; | 99US-0138847P. | |
| PR | 14-JUN-1999; | 99US-0139119P. | |
| PR | 16-JUN-1999; | 99US-0139452P. | |
| PR | 16-JUN-1999; | 99US-0139453P. | |
| PR | 17-JUN-1999; | 99US-0139492P. | |
| PR | 18-JUN-1999; | 99US-0139454P. | |
| PR | 18-JUN-1999; | 99US-0139455P. | |
| PR | 18-JUN-1999; | 99US-0139456P. | |
| PR | 18-JUN-1999; | 99US-0139457P. | |
| PR | 18-JUN-1999; | 99US-0139458P. | |
| PR | 18-JUN-1999; | 99US-0139459P. | |
| PR | 18-JUN-1999; | 99US-0139460P. | |
| PR | 18-JUN-1999; | 99US-0139461P. | |
| PR | 18-JUN-1999; | 99US-0139462P. | |
| PR | 18-JUN-1999; | 99US-0139463P. | |
| PR | 18-JUN-1999; | 99US-0139750P. | |

```

PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139889P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156566P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161923P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 24.9%; Score 182.6; DB 3; Length 1005;
Best Local Similarity 70.0%; Pred. No. 3e+42;
Matches 296; Conservative 0; Mismatches 109; Indels 18; Gaps 3;

QY 270 CTGATCTTCTAATCCAAACGACAGACAGAAACAACTCTGATGATGATTCGAGAAAT 329
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 CTGATTTTCCAGATCAAACTTCGAAAAGAGAAAAGCCTTCTTAATGATTCGTGAGTT 558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 330 CTCAAACTTACAAAAGATATTTAATCATTCAGACAGATTCGAAACCAAGCAT 389
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 CTCGAGTCTACAAAACGACATCATTAACATTTCAAGAACATTTCTCTCAACCAACAT 618
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 390 GGAAGTCGAGAA---AACAACGGAAGAAACATGAGCTTCGAAGTT-----CGCC 437
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 619 GGAAGTGAAGAGATGATTAACGGAAGAAACAGAGTTTGTGTGTTCAAGTCCAGCTCC 678
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 438 GGAAGAAAGCAAGCAAGATTACGCGGAGTTTCTCGCGCAGAA-----AAGATGAT 491
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 679 GATTATGCGGAAGCAAGAAATGTGTTCCGACATTCCTCGCAGAACATCCCAAGAGGCT 738
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 492 GATCCGCGGAATATTATTCAGTTAGATGCTGTTCTATTGTTGTTGTTGTTAAGGATC 551
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 739 AATATACGACGATTACTTCACTTGAATCAATGGTTGCTGTGCTCAACAGCATC 798
 Qy 552 TCTGTATATCTTCCGTTAGTTTGGCCGCTTACTCCGCTCCGTTTATNGCTATTT 611
 Db 799 TCTCTTATCTTACCTTGAATCTTCCACCATTCCTCCTCTTATATGCTCTTTT 858
 Qy 612 GGTTCCTATGGGATTAATGGTTTACTCGTGTCTTACCTTCAATGCTTCTCAATTC 671
 Db 859 GATTCCTATGGGATTAATGGTTTCTTATGTTCTTCAATGCTTCTTCAATTC 918
 Qy 672 TAA 674
 Db 919 CAA 921

RESULT 3
 AB213462
 ID AB213462 standard; DNA, 336 BP.
 AC AB213462;
 DT 21-JAN-2003 (first entry)
 XX
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1267.
 XX
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200216655-A2.
 XX
 PD 28-FEB-2002.
 PF 24-AUG-2001; 2001WO-US026685.
 XX
 PR 24-AUG-2000; 2000US-0227866P.
 XX
 PR 26-JAN-2001; 2001US-0264647P.
 XX
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PI Harper JF, Krepe J, Wang X, Zhu T;
 PI
 DR WPI; 2002-304127/34.
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 144; SEQ ID NO 1267; 577pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 XX Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;
 SQ

Query Match 18.1%; Score 132.6; DB 6; Length 336;
 Best Local Similarity 71.3%; Pred. No. 6,1e-28;
 Matches 209; Conservative 0; Mismatches 69; Indels 15; Gaps 2;
 Qy 397 GGAAGAAACACCGGAAAAACATGAGCTTTCGAAGTT-----CGCCGAGAAAAAGC 447
 Db 13 GGAGATCATACCGGAAAAACACGAGTTTCTGGTTCAGCTCCAGCTCGAATTATGGGG 72

Qy 448 AACCAAGATTACGGGAGATTCTCGGCGCAGAA-----AAGGATGATGATCCGGCG 501
 Db 73 AACCAAGATTGTTTGGACATTTGTGTGCGAAGAACGCAAGAGCTAATATACGG 132
 Qy 502 AATTATTCAGTTTAAAGTCTGTCTCTTATTTGTTGTTAAAGGCAATCTGTATA 561
 Db 133 AGTTACTTCAGTTTAAATCAATGATGTTGTTGTTGTTCAACGATCTCTTGATC 192
 Qy 562 CTTCGTTTATGTTTGGCGCGTTACTCCGCTCCGTTTATGCTGATTTGTTCCATT 621
 Db 193 TTACGTTTATGTTTCTTCCACATTCCTCCTCTTATGCTGTTGATTCCTATT 252

Qy 622 GGATTAATGTTTACTTCGCTGTTGCTTCAATGCTTCAATGCTTCAATGCTTCA 674
 Db 253 GGATTAATGTTTACTTCGCTGTTGCTTCAATGCTTCAATGCTTCAATGCTTCA 305

RESULT 4
 ADA67786
 ID ADA67786 standard; DNA, 336 BP.
 AC ADA67786;
 DT 20-NOV-2003 (first entry)
 XX
 DE Arabidopsis thaliana gene, SEQ ID 23.
 XX
 KW Plant; bacterial infection; fungal infection; viral infection; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2003000898-A1.
 XX
 PD 03-JAN-2003.
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
 PI
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 6; SEQ ID NO 23; 899pp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 XX Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;
 SQ

Query Match 18.1%; Score 132.6; DB 8; Length 336;
 Best Local Similarity 71.3%; Pred. No. 6,1e-28;
 Matches 209; Conservative 0; Mismatches 69; Indels 15; Gaps 2;
 Qy 397 GGAAGAAACACCGGAAAAACATGAGCTTTCGAAGTT-----CGCCGAGAAAAAGC 447
 Db 13 GGAGATCATACCGGAAAAACACGAGTTTCTGGTTCAGCTCCAGCTCGAATTATGGGG 72

| | | |
|----|-------------|---------------|
| PR | 13-AUG-1999 | 99US-0147038P |
| PR | 04-AUG-1999 | 99US-0147202P |
| PR | 04-AUG-1999 | 99US-0147302P |
| PR | 05-AUG-1999 | 99US-0147719P |
| PR | 05-AUG-1999 | 99US-0147260P |
| PR | 06-AUG-1999 | 99US-0147303P |
| PR | 06-AUG-1999 | 99US-0147416P |
| PR | 09-AUG-1999 | 99US-0147493P |
| PR | 09-AUG-1999 | 99US-0147955P |
| PR | 10-AUG-1999 | 99US-0148171P |
| PR | 11-AUG-1999 | 99US-0148319P |
| PR | 12-AUG-1999 | 99US-0148341P |
| PR | 13-AUG-1999 | 99US-0148565P |
| PR | 13-AUG-1999 | 99US-0148664P |
| PR | 16-AUG-1999 | 99US-0149368P |
| PR | 17-AUG-1999 | 99US-0149417P |
| PR | 18-AUG-1999 | 99US-0149415P |
| PR | 25-AUG-1999 | 99US-0150565P |
| PR | 26-AUG-1999 | 99US-0150884P |
| PR | 20-AUG-1999 | 99US-0149722P |
| PR | 20-AUG-1999 | 99US-0149723P |
| PR | 27-AUG-1999 | 99US-0149929P |
| PR | 27-AUG-1999 | 99US-0149929P |
| PR | 27-AUG-1999 | 99US-0149929P |
| PR | 30-AUG-1999 | 99US-0151303P |
| PR | 31-AUG-1999 | 99US-0151438P |
| PR | 13-SEP-1999 | 99US-0153758P |
| PR | 15-SEP-1999 | 99US-0154018P |
| PR | 16-SEP-1999 | 99US-0154039P |
| PR | 20-SEP-1999 | 99US-0154779P |
| PR | 22-SEP-1999 | 99US-0155139P |
| PR | 23-SEP-1999 | 99US-0155466P |
| PR | 24-SEP-1999 | 99US-0155659P |
| PR | 28-SEP-1999 | 99US-0156458P |
| PR | 29-SEP-1999 | 99US-0156596P |
| PR | 04-OCT-1999 | 99US-0157117P |
| PR | 05-OCT-1999 | 99US-0157753P |
| PR | 06-OCT-1999 | 99US-0157865P |
| PR | 07-OCT-1999 | 99US-0158022P |
| PR | 08-OCT-1999 | 99US-0158212P |
| PR | 12-OCT-1999 | 99US-0158369P |
| PR | 13-OCT-1999 | 99US-0159283P |
| PR | 13-OCT-1999 | 99US-0159294P |
| PR | 14-OCT-1999 | 99US-0159295P |
| PR | 14-OCT-1999 | 99US-0159331P |
| PR | 14-OCT-1999 | 99US-0159310P |
| PR | 14-OCT-1999 | 99US-0159637P |
| PR | 14-OCT-1999 | 99US-0159638P |
| PR | 18-OCT-1999 | 99US-0159584P |
| PR | 21-OCT-1999 | 99US-0160741P |
| PR | 21-OCT-1999 | 99US-0160767P |
| PR | 21-OCT-1999 | 99US-0160768P |
| PR | 21-OCT-1999 | 99US-0160770P |
| PR | 21-OCT-1999 | 99US-0160814P |
| PR | 21-OCT-1999 | 99US-0160815P |
| PR | 22-OCT-1999 | 99US-0160960P |
| PR | 22-OCT-1999 | 99US-0160981P |
| PR | 22-OCT-1999 | 99US-0160989P |
| PR | 23-OCT-1999 | 99US-0161040P |
| PR | 25-OCT-1999 | 99US-0161405P |
| PR | 25-OCT-1999 | 99US-0161406P |
| PR | 26-OCT-1999 | 99US-0161359P |
| PR | 26-OCT-1999 | 99US-0161360P |
| PR | 26-OCT-1999 | 99US-0161361P |
| PR | 28-OCT-1999 | 99US-0161920P |
| PR | 28-OCT-1999 | 99US-0161992P |

| | | |
|----|---------------------------|--|
| PR | 28-OCT-1999; | 99US-0161993P. |
| PR | 29-OCT-1999; | 99US-0162142P. |
| | Query Match | 15.0%; Score 109.8; DB 3; Length 650; |
| | Best Local Similarity | 74.6%; Pred. No. 3.3e-21; |
| | Matches 138; Conservative | 0; Mismatches 47; Indels 0; Gaps 0; |
| OY | 490 | ATGATCCCGGAGATTATTTCAAGTTAAGTCTCTGTCCATTAGTGTGCTTAAAGCA 549 |
| DB | 402 | ATTATCAGAGTTGATCTGTGTTTTTAAGCTCAATAGTGTGCTGTGCTCACACA 461 |
| OY | 550 | TCTCGTTAATACTCCGTTAGTTTTCGCGCGATTACCTCGGCGCTCGTTATAGCTGCTA 609 |
| DB | 462 | TCTCTCTTGATCTTACCGTTGATTTCTTCACACATTGCTCTCTCCCTTTTATAGCTGCT 521 |
| OY | 610 | TTGGTTCCTCATTCGAGTTATAGGTTTACTCGTCGTTCTTGCCCTTCAGGCTTTCTTCAT 669 |
| DB | 522 | TTCATTTCTATATGGAGTTATAGTGTTCATTATAGTTTCTTGCTTCATGCTCTTCTTCAT 581 |
| OY | 670 | TCTAA 674 |
| DB | 582 | TCCAA 586 |

RESULT 6
 ACT31088/c
 ID ACT31088 standard; cDNA, 298 BP.
 XX
 AC ACT31088;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress response related polynucleotide SEQ ID NO:9651.
 XX
 KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 XX agriculture.
 XX
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 XX 24-AUG-2001; 2001US-0314662P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Krups J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
 PI Moughamer T, Frowart N, Rieke D, Zhu T;
 XX
 DR WPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 PT
 PS Disclosure; SEQ ID NO 9651; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

KM growth rate; cell cycle pathway; disease resistance;
 KM galactomannan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content; gene; ss.
 XX Unidentified.
 OS
 PN US2004034888-A1.
 PD
 PD 19-FEB-2004.
 XX
 XX 28-APR-2003; 2003US-00425114.
 PF
 PF 06-MAY-1999; 99US-00304517.
 PR 03-NOV-2001; 2001US-00985678.
 XX
 XX (LIU/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 DR WPI; 2004-180133/17.
 XX
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 XX Claim 1; SEQ ID NO 24068; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 XX Sequence 665 BP; 111 A; 228 C; 205 G; 121 T; 0 U; 0 Other;
 QY
 QY Query Match 8.3%; Score 60.6; DB 13; Length 665;
 Best Local Similarity 64.7%; Pred. No. 6.4e-07;
 Matches 90; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 527 TCCATATGTTGGTGTAAACGCGATCTCTGTTAATACCTTCGTTTGGCGCGCTTAC 586
 Db 521 TGTGCTGTGTGGGCTACCGCGTCCGCGTGGAGATCTCCCGCTGCTCCGCCGCTGC 580
 QY 587 CTCGCGCTCCGTTTATGCTGCTATGTTCCCATTTGGATTATGTTTACTCTGCTGC 646
 Db 581 CGCGCGCGCGCTGCTGCTATGCTGCTGCTCCCGTGGCAATCTCTCTGCTGCTGCTGC 640
 QY 647 TTGCTTTCATGCTTCTTCT 665
 Db 641 TGGGCTTCATGCTCCACGCTC 659
 RESULT 11

ADT18995
 ID ADT18995 standard; cDNA; 1055 BP.
 XX
 XX ADT18995;
 AC
 XX
 XX 13-JAN-2005 (first entry)
 DT
 XX
 XX Plant cDNA, Seq ID 4321.
 DE
 XX
 XX Plant; ss; gene; transgenic; cold tolerance; growth rate;
 KM drought tolerance; disease resistance; galactomannan production;
 KM plant growth regulator; heat tolerance; herbicide tolerance;
 KM lignin production; extreme osmotic condition tolerance;
 KM pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KM seed protein yield.
 XX
 XX Viridiplantae.
 OS
 OS US2004216190-A1.
 PN
 XX
 XX 28-OCT-2004.
 PD
 XX
 XX 18-DEC-2003; 2003US-00739930.
 PF
 XX
 XX 28-APR-2003; 2003US-00424599.
 PR 28-APR-2003; 2003US-00425115.
 XX
 XX (KOVA/) KOVALIC D K.
 PA
 XX
 XX Kovalic DK;
 PI
 XX
 XX WPI; 2004-757369/74.
 DR
 XX
 XX New recombinant DNA constructs useful in the field of biotechnology and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 PT
 XX
 PS Claim 1; SEQ ID NO 4321; 14pp; English.
 PS
 XX
 XX The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed
 CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biotechnology and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did

Query Match 7.7%; Score 56; DB 6; Length 2000;
 Best Local Similarity 68.8%; Pred. No. 2.2e-05;
 Matches 77; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 270 CTGATCTTCTTAATCCAAACGACAGCAACCAACTCTGTAGATGATTCGAGAAAT 329
 |||||
 DB 1883 CTGATTTCCAGATCAACTTCGAAAAGAGAAAAGCCCTTTTAATGATTCGTGAGTT 1942
 |||||

QY 330 CTCGAACCTTCAAAAAGATATTTAATCAATTCAAGACAGATTATTCGAACAAC 381
 |||||
 DB 1943 CTCGAGCTCAAAAACGACATCATTAACATTCAGAACATTATCTCTCAAC 1994
 |||||

RESULT 14
 ADA69206
 ID ADA69206 standard; DNA; 2000 BP.
 XX ADA69206;
 AC
 XX 20-NOV-2003 (first entry)
 DT
 XX Arabidopsis thaliana gene, SEQ ID 2529.
 DE
 XX Plant; bacterial infection; fungal infection; viral infection; gene; ds.
 KM
 XX Arabidopsis thaliana.
 OS
 XX Arabidopsis thaliana.
 PN
 XX MO2003000898-A1.
 PD
 XX 03-JAN-2003.
 PF
 XX 22-JUN-2001; 2001MO-IB001105.
 PR
 XX 22-JUN-2001; 2001MO-IB001105.
 XX
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katsagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 PS
 XX Claim 27; SEQ ID NO 2529; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 CC
 SO Sequence 2000 BP; 710 A; 365 C; 235 G; 690 T; 0 U; 0 Other;

Query Match 7.7%; Score 56; DB 8; Length 2000;
 Best Local Similarity 68.8%; Pred. No. 2.2e-05;
 Matches 77; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 270 CTGATCTTCTTAATCCAAACGACAGCAACCAACTCTGTAGATGATTCGAGAAAT 329
 |||||
 DB 1883 CTGATTTCCAGATCAACTTCGAAAAGAGAAAAGCCCTTTTAATGATTCGTGAGTT 1942
 |||||

QY 330 CTCGAACCTTCAAAAAGATATTTAATCAATTCAAGACAGATTATTCGAACAAC 381
 |||||

DB 1943 CTCGAGCTCAAAAACGACATCATTAACATTCGAAGAATTTCTCTCAAC 1994
 |||||

RESULT 15
 ID ADR64525
 ID ADR64525 standard; cDNA; 366 BP.
 XX ADR64525;
 AC
 XX 02-DEC-2004 (first entry)
 DT
 XX Cotton cDNA sequence, SEQ ID 5306.
 DE
 XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
 KM drought tolerance; plant disease resistance; galactomannan; lignin;
 KM plant growth regulator; heat tolerance; herbicide tolerance;
 KM homologous recombination; extreme osmotic condition tolerance;
 KM pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
 KM stress resistance.
 XX
 XX Gossypium hirsutum.
 OS
 XX US2004181830-A1.
 PN
 XX 16-SEP-2004.
 PD
 XX 29-JAN-2004; 2004US-00767795.
 PE
 XX 07-MAY-2001; 2001US-00849529.
 PR 12-DEC-2001; 2001US-00021323.
 XX
 XX (KOVA/) KOVALIC D K.
 PA (ZHOU/) ZHOU Y.
 PA (CAOY/) CAO Y.
 XX
 XX Kovalic DK, Zhou Y, Cao Y;
 PI
 XX WPI; 2004-667718/65.
 DR
 XX
 PT New recombinant nucleic acid molecules and polypeptides from Gossypium
 PT hirsutum, useful for producing plants with improved biological
 PT characteristics (e.g. improved plant cold or drought tolerance).
 PT
 XX
 XX Claim 1; SEQ ID NO 5306; 14dp; English.
 PS
 XX The invention relates to a recombinant polynucleotide comprising any of
 CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
 CC Also a recombinant polypeptide comprising any of the 58798 amino acid
 CC sequences mentioned in the specification and producing a plant having an
 CC improved property. Producing a plant having an improved property
 CC comprises transforming a plant with a recombinant construct comprising a
 CC promoter region functional in a plant cell operably joined to a
 CC polynucleotide comprising a coding sequence for a polypeptide associated
 CC with the property, and growing the transformed plant. The polypeptide is
 CC useful for improving plant cold tolerance, manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, improving plant
 CC drought tolerance, providing increased resistance to plant disease,
 CC producing galactomannan (or lignin or plant growth regulators), improving
 CC plant heat tolerance, improving plant tolerance to herbicides, increasing
 CC the rate of homologous recombination in plants, improving plant tolerance
 CC to extreme osmotic conditions or to pathogens or pests, improving yield
 CC by modification of photosynthesis, modifying seed oil or protein yield
 CC and/or content, improving yield by modification of carbohydrate, nitrogen
 CC or phosphorus use and/or uptake, or improving yield by providing improved
 CC plant growth and development under at least one stress condition. The
 CC polynucleotide and polypeptide may also be used in recombinant DNA
 CC constructs, in physical arrays of molecules, as plant breeding markers,
 CC or in computer-based storage and analysis systems. The present sequence
 CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=20040181830. However only 6585
 CC polynucleotide sequences were available, the remaining 52213

CC polynucleotides and all 58798 protein sequences were not present.
XX
SQ Sequence 366 BP; 65 A; 112 C; 71 G; 118 T; 0 U; 0 Other;

XX Sequence 366 BP; 65 A; 112 C; 71 G; 118 T; 0 U; 0 Other;
SQ

| | | | | | | | |
|-----------------------|-------|--------------|--------|------------|----|--------|-----|
| Query Match | 6.3% | Score | 45.8 | DB | 13 | Length | 366 |
| Best Local Similarity | 58.4% | Pred. No. | 0.0097 | | | | |
| Matches | 80 | Conservative | 0 | Mismatches | 57 | Indels | 0 |
| | | | | | | Gaps | 0 |

| | | | |
|----|-----|--|-----|
| QY | 527 | TCCTATATGGTTGGTCTAAAGGCATCTGTATATCTCCGATGTTTGGCGCGGTAC | 586 |
| Db | 138 | TGCTTTGGCCGTGTTCGTGTGCTCTCTTTGGTGCTGCTCTGAGTGTGGCCAAATGC | 197 |
| QY | 587 | CTCCGCCTCCGTTTATGCTGCTATTGTTCCCATGGAGTTATGTTTACTGTGCTTC | 646 |
| Db | 198 | CGCCTCCGCCCCCTTCGCTCTTCTTAATCCCAATTGATGATGATGCTGCTCTCTTTCC | 257 |
| QY | 647 | TTGACCTGATGCCTCT | 663 |
| Db | 258 | TGGCATTTTCAACGCTCT | 274 |

Search completed: January 30, 2006, 04:28:39
Job time : 514 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:56:47 ; Search time 185 Seconds

(without alignments)

7033.379 Million cell updates/sec

Title: US-10-715-129-1

Sequence: 1 tttgtctccctccattccctca.....gaaatttatatgatgat 732

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PC/US_COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 49.6 | 6.8 | 7218 | 2 | US-08-232-463-14 |
| 2 | 40.6 | 5.5 | 142783 | 3 | US-09-949-016-15127 |
| 3 | 40 | 5.5 | 832 | 3 | US-09-621-976-2813 |
| 4 | 39.8 | 5.4 | 1141 | 3 | US-09-806-708B-22 |
| 5 | 38.4 | 5.2 | 2651 | 3 | US-09-976-594-1050 |
| 6 | 38 | 5.2 | 197131 | 3 | US-09-949-016-12675 |
| 7 | 38 | 5.2 | 197132 | 3 | US-09-949-016-17170 |
| 8 | 37.6 | 5.1 | 64923 | 3 | US-09-949-016-12901 |
| 9 | 37.6 | 5.1 | 65902 | 3 | US-09-949-002-609 |
| 10 | 37.4 | 5.1 | 1141 | 3 | US-09-806-708B-22 |
| 11 | 37.4 | 5.1 | 64593 | 3 | US-09-949-016-16654 |
| 12 | 37.4 | 5.1 | 64593 | 3 | US-09-949-016-16655 |
| 13 | 37.4 | 5.1 | 64593 | 3 | US-09-949-016-16656 |
| 14 | 37.4 | 5.1 | 99498 | 3 | US-09-949-016-12621 |
| 15 | 37.4 | 5.1 | 146428 | 3 | US-09-949-016-12620 |
| 16 | 37.4 | 5.1 | 146438 | 3 | US-09-949-016-12081 |
| 17 | 37.2 | 5.1 | 601 | 3 | US-09-949-016-18101 |
| 18 | 37.2 | 5.1 | 601 | 3 | US-09-949-016-18102 |
| 19 | 37.2 | 5.1 | 601 | 3 | US-09-949-016-18103 |
| 20 | 37.2 | 5.1 | 601 | 3 | US-09-949-016-51599 |
| 21 | 37.2 | 5.1 | 601 | 3 | US-09-949-016-51600 |
| 22 | 37.2 | 5.1 | 601 | 3 | US-09-949-016-51601 |
| 23 | 37.2 | 5.1 | 63658 | 3 | US-09-949-016-13238 |
| 24 | 37.2 | 5.1 | 64489 | 3 | US-09-949-016-11766 |

| | | | | | | |
|------|------|-----|--------|---|----------------------|-------------------|
| C 25 | 36.8 | 5.0 | 260286 | 3 | US-09-949-016-17037 | Sequence 17037, A |
| C 26 | 36.8 | 5.0 | 260289 | 3 | US-09-949-016-12106 | Sequence 12106, A |
| C 27 | 36.6 | 5.0 | 145287 | 3 | US-09-949-016-13530 | Sequence 13530, A |
| C 28 | 36.6 | 5.0 | 145287 | 3 | US-09-949-016-13531 | Sequence 13531, A |
| C 29 | 36.2 | 4.9 | 9821 | 3 | US-08-956-171B-470 | Sequence 470, App |
| C 30 | 36.2 | 4.9 | 9821 | 3 | US-08-781-986A-470 | Sequence 470, App |
| C 31 | 36.2 | 4.9 | 43095 | 3 | US-09-676-519-17 | Sequence 17, Appl |
| C 32 | 36 | 4.9 | 157644 | 3 | US-09-949-016-16179 | Sequence 16179, A |
| C 33 | 36 | 4.9 | 157644 | 3 | US-09-949-016-16180 | Sequence 16180, A |
| C 34 | 35 | 4.8 | 300 | 3 | US-09-248-796A-12919 | Sequence 12919, A |
| C 35 | 35 | 4.8 | 135667 | 3 | US-09-949-016-15051 | Sequence 15051, A |
| C 36 | 35 | 4.8 | 152486 | 3 | US-09-949-016-12869 | Sequence 12869, A |
| C 37 | 35 | 4.8 | 253345 | 3 | US-09-949-016-12856 | Sequence 12856, A |
| C 38 | 35 | 4.8 | 253364 | 3 | US-09-949-016-13639 | Sequence 13639, A |
| C 39 | 35 | 4.8 | 451924 | 3 | US-09-949-016-12896 | Sequence 12896, A |
| C 40 | 35 | 4.8 | 451925 | 3 | US-09-949-016-17305 | Sequence 17305, A |
| C 41 | 34.8 | 4.8 | 7218 | 2 | US-08-232-463-14 | Sequence 14, Appl |
| C 42 | 34.8 | 4.8 | 218940 | 3 | US-09-949-016-17539 | Sequence 17539, A |
| C 43 | 34.6 | 4.7 | 920 | 3 | US-09-270-767-26471 | Sequence 26471, A |
| C 44 | 34.6 | 4.7 | 1524 | 3 | US-09-270-767-10971 | Sequence 10971, A |
| C 45 | 34.6 | 4.7 | 2169 | 3 | US-09-434-408-3 | Sequence 3, Appl1 |

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-F18

US-08-232-463-14

Query Match
Best Local Similarity 6.8%; Score 49.6; DB 2; Length 7218;
Matches 19; Conservative 160; Mismatches 109; Indels 0; Gaps 0;

QY 1 TTGCTCTCATTTCCCTACTAGTACTGTTTCACAGAGTTCTGATCCACCAAC 60
D 1183 YY 1242
QY 61 CAAATACCAAGCTTCACAACTCTCACTCAAGCTTCTCTTAATCTGAATCG 120
D 1243 YY 1302
QY 121 TTGAGTAACTCGGATTTGTTCTGATCTGTTTGAATGAGGCACTCTATT 180
D 1303 YY 1362
QY 181 TGCTCGAATCTTCACCAATGCTTCGATCAAGCTGATGTTACAGTGCCTAA 240
D 1363 YY 1422
QY 241 AGATCAATCTTTGACCAAAATTTGTCAGTCACTTCTTAATCCAA 288
D 1423 YYYYYYYYYYGTACCAAAATCTTCTATCTCTTAATCTTGCAATA 1470

RESULT 2

US-09-949-016-15127
Sequence 15127, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 15127
LENGTH: 142783
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(142783)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15127

Query Match
Best Local Similarity 5.5%; Score 40.6; DB 3; Length 142783;

Matches 85; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 272 GATCTCTAAATCCAAACAGACACAGCAAAACAACTCTGTAGATGTTGCAAAATCT 331
D 140872 GACTTCGAAAGCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 140931
QY 332 CAAACTTACAAAAGATTTTAAACATTCAAGACAGTTATTCGAACACCGACTCATG 391
D 140932 CACACACACAAAAGAGAAAACCTAACAGTTAATTTTAAAAAAGAACAAATA 140991
QY 392 ACGTCGAGAAACAAACCGGAAAAACATGAGCTTTCGA 430
D 140992 TAAGGAAACAAAGACAAACCAACGGAATGTGGAA 141030

RESULT 3

US-09-621-976-2813
Sequence 2813, Application US/09621976
Patent No. 663963

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621, 976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 5.5%; Score 40; DB 3; Length 832;
Matches 43; Conservative 174; Mismatches 179; Indels 0; Gaps 0;

QY 178 TTTTGTCTGAATCTTCCACCAATTTGCTTCGATCAAGCTGATGTTAACAGAGTGGCC 237
D 5 WMYRTTYWYACWTKMWSWYMMYKMYKTYMRRRKKKAKWKKYKTYWYMYAM 64
QY 238 TAAAGATGATCTTTGAGCAAAATTTTGCAGTCTCTTAATCCAAACAGACACA 297
D 65 WGYKKKAMKRTTKKKKKGGYMMYTWGRBSYMAATTTWGYAYRSMYMKRCH 124
QY 298 GCAAAACAACTCTGTAGATTTGCAAGAAATCTCAAACTTACAAAAGATATTATTAAC 357
D 125 KKAAYRKTCYSSKGMWTKMKWKAATWMMKKTYMAATRYMMWCMWKRRBASWYC 184
QY 358 ATTCAAGACAGTTATTTGCAACCAAGCATGATGACCTCGAAGAAACACCGAATAAC 417
D 185 WMMKARKMYTKRSYSASARSAKCYSCSWGASWKMWMMRWGATGAGMKAMR 244
QY 418 ATAGCTTTCGAGTTCGCCGCGAAGAAAGCAAGAGTTAGCGGAGTTCCTCGCG 477
D 245 ASCMRRKTAGSKSTYSKMMCTTSWKYCTKARTGYCYRKGGMMGRGRWASKK 304
QY 478 CAGAAAAGATGATATCCCGGCAATTTATTCAGTTAGATGCTCTGTTCTATTGGTT 537
D 305 YMKRMMWCMWABMYRSTGTBRABMMWRMYTMMWMMWKKYAWABAAWMMWABRAC 364
QY 538 GGTCTAAGGCACTCTGTTAATACTTCGTTGTT 573
D 365 AAATATTAATTAATTAATGATCAATCTTGTACTTT 400

RESULT 4

US-09-806-708B-22
Sequence 22, Application US/09806708B
Patent No. 6784342

GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806, 708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA

RESULT 7

US-09-949-016-17170/c
; Sequence 17170, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17170

LENGTH: 197132

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(197132)

OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17170

Query Match 5.2%; Score 38; DB 3; Length 197132;
Best Local Similarity 54.2%; Pred. No. 3.6;
Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 237 CTAAGATCGATCTTGAGCAAAATTTGTCTACTGATCTTCTTAAACCAACGACAC 236

DB 44714 CTTAAGCCCAAGATTTTAAAGCAATAGTGAAGCTATGATGACGACGACGACGCTTA 44655

QY 297 AGCAAAACAACCTCTGTAGATGATTCGAGAAATCTCAAACTTACAAAGATTTATTA 356

DB 44654 GCGACACGACGACGACGCTGCTTTAAATAAATAAATAAATAAATAAATAAATAA 44595

QY 357 CATTCAAGACAGTTATTGCAAC 378

DB 44594 GAATGAGAAAAGTTTATGATC 44573

RESULT 8

US-09-949-016-12901/c
; Sequence 12901, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12901

LENGTH: 64923

TYPE: DNA

ORGANISM: Human

US-09-949-016-12901

Query Match 5.1%; Score 37.6; DB 3; Length 64923;

Best Local Similarity 46.3%; Pred. No. 2.6;
Matches 124; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 219 ATTGTTAACCACTGTCCTTAAAGATCATCTTTGAGCAAAATTTGTCTACTGATCTTC 278

DB 53049 ATTGTATTTCCAAAGAAAGATATAGATGTGAGAAATGATTAATTTTAAAGTAAATGTC 52990

QY 279 TAATATCAAAACGACACGACAAACCACTCTGTAGATGATTCGAGAAATCTCAACTT 338

DB 52989 TAGAGTTTCCAGACTTGAGAGAAATCCCACTCTCAGATTCAGAAATCTCACTAAT 52930

QY 339 ACAAAGATATATTAATTAACATTCAGACATTAATTCGAACCAACGATCATGACGTCG 398

DB 52929 CCAAGCTATACATTAATTAATTAAGAAATTAATCATCTAGACATTAAGAGGAAACT 52870

QY 399 AAGAAACAACCGGAAAAACATGAGCTTTGGAATGTCGCGGAGAAAAACGACGAGATT 458

DB 52869 TTAAGCAACATGATACAAAACTTTAAAGTAGCAGAGAGAAAGAAACTAGACTG 52810

QY 459 ACGGCGAGTTTCTCGGCGAGAAAAAG 486

DB 52809 ACAGCAGATTTCTTACGACAGTAGAAG 52782

RESULT 9

US-09-949-002-609/c
; Sequence 609, Application US/09949002
; Patent No. 6900016

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1000790

CURRENT APPLICATION NUMBER: US/09/949,002

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: 60/231,401

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 10823

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 609

LENGTH: 65902

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(65902)

OTHER INFORMATION: n = A,T,C or G

US-09-949-002-609

Query Match 5.1%; Score 37.6; DB 3; Length 65902;
Best Local Similarity 52.6%; Pred. No. 2.6;
Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 278 CTAATCCAAACGACACGACAAACCACTCTGTAGATGATTCGAGAAATCTCAACT 337

DB 40317 CTTAAACCAAGCAAAACCACTACGACTCCCTTAAAGATTAACAGAAATGAAACA 40258

QY 338 TACAAAGATATTAATTAACATTCAGACATTAATTCGAACCAACGATCATGACGTCG 397

DB 40257 AGAACAGAGATGTTATTAATAAACAAGAGCTCTTGATTAATAAAGCATGACATG 40198

QY 398 GAAGAAACAACCGGAAAAACATGAGCTTTGCAAGTT 433

DB 40197 AAAAGTAAACTTAATTAACAAGCTGGGTTGGAGAT 40162

RESULT 10

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342

GENERAL INFORMATION:

APPLICANT: The University of British Columbia

US-09-806-708B-22/c


```

RESULT 13
US-09-949-016-16656/c
: Sequence 16656, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 16656
: LENGTH: 64593
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(64593)
: OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16656

Query Match          5.1%; Score 37.4; DB 3; Length 64593;
Best Local Similarity 51.5%; Pred. No. 3;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

OY      550  TCTCTGTAATATCTCCGTTAGTTTGGCCGCCGTAACTCCGCTCGTTTANGTCTGA 609
Db      14588 TCACGAAATTTCACCACTGTTGTTTGGCTTCACTTGACCTCTTAAGTTATGTGTA 14529

OY      610  TTGATGCCATGGGAGTTATGTTTACTCGTCTGCTTCCGCTTCAAGCCTTCTGAT 669
Db      14528 TTACTTCGATGGGACATCTGTTTGACAGGTGAGTGCATCTTGGCTTATTTCTT 14469

OY      670  TCTAATGCTAATACAGATGTAACTTGCAATTTGCATGTAAATCTGAAA 716
Db      14468 ACAATGGCCCAATTCTGTTTACTCTGCTAATTAATGAATAATGGAAA 14422

RESULT 14
US-09-949-016-12621/c
: Sequence 12621, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12621
: LENGTH: 99498
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(99498)

```

```

US-09-949-016-12621
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12621

Query Match          5.1%; Score 37.4; DB 3; Length 99498;
Best Local Similarity 51.5%; Pred. No. 3.8;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 550 TCTCTGTTAAACTCCGCTGTTAGTTTGGCCGCGCTTACCTCCGCGCTCCGTTTAAAGCTGCTA 609
DB 49493 TCACGTAAATTACCACTGTGTGTTTGGCTTCATCTGACCTCTTAAGTTAATGTAATA 49434

QY 610 TTGGTTCCATTGGGATTAATGATTTTACCTGCTGCTTCCCTCATGAGCTTCTTCTCAT 669
DB 49433 TTTACTTCATGAGGAGATCTGTGTTGGCAGGTTGAAGTGCATCTTGTGCTTATTTCCT 49374

QY 670 TCTAATGCTAATACAGATGTAACTTGCATTTCAATGTAATCTGAAA 716
DB 49373 ACAATGGCCCAATTCTGTTTACCTGTATATTAATTAATGAGATA 49327

RESULT 15
US-09-949-016-12620/c
; Sequence 12620, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12620
; LENGTH: 146428
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc.feature
; LOCATION: (1) ..(146428)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12620

Query Match          5.1%; Score 37.4; DB 3; Length 146428;
Best Local Similarity 51.5%; Pred. No. 4.7;
Matches 86; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 550 TCTCTGTTAAACTCCGCTGTTAGTTTGGCCGCGCTTACCTCCGCGCTCCGTTTAAAGCTGCTA 609
DB 48423 TCACGTAAATTACCACTGTGTGTTTGGCTTCATCTGACCTCTTAAGTTAATGTAATA 48364

QY 610 TTGGTTCCATTGGGATTAATGATTTTACCTGCTGCTTCCCTCATGAGCTTCTTCTCAT 669
DB 48363 TTTACTTCATGAGGAGATCTGTGTTGGCAGGTTGAAGTGCATCTTGTGCTTATTTCCT 48304

QY 670 TCTAATGCTAATACAGATGTAACTTGCATTTCAATGTAATCTGAAA 716
DB 48303 ACAATGGCCCAATTCTGTTTACCTGTATATTAATTAATGAGATA 48257

```

Search completed: January 30, 2006, 06:39:09
Job time : 189 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 30, 2006, 08:30:29 ; Search time 2493 Seconds
(without alignments) 2416.930 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533
Sequence: 1 MDVGRNNRKNMSFRSSPEKS.....LAFWPSHSNANTDVTCNFM 106

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 2842172553 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2.1/USPRO.spool.p/US10715129/runat.27012006.154744.14566/app.query.fasta.1.263
-DB=genmbl -GPM=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORHEX=HEXSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10715129 -OCGN 1.1 7415 -ORUNAT 27012006.154744.14566 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb Da:*
2: gb In:*
3: gb Env:*
4: gb Om:*
5: gb Ov:*
6: gb Pat:*
7: gb Ph:*
8: gb Pr:*
9: gb Ro:*
10: gb Sts:*
11: gb Sy:*
12: gb Un:*
13: gb Vi:*
14: gb Htg:*
15: gb Pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|-------|--------------------|
| 1 | 533 | 100.0 | 393 | BT011724 Arabidops |
| 2 | 533 | 100.0 | 732 | AY305869 Arabidops |
| 3 | 533 | 100.0 | 747 | AK176640 Arabidops |

4 533 100.0 100906 15 ATP24G16
5 327 61.4 336 6 AX506572
6 327 61.4 336 6 AX651235
7 327 61.4 439 15 BT000930
8 327 61.4 904 15 AY080817
9 327 61.4 109741 15 AC004005
10 224.5 42.1 1027 15 AC061481
11 224.5 42.1 1137 15 AK060546
12 224.5 42.1 110000 15 AP008217_078
13 224.5 42.1 180999 15 AC137924
14 209 39.2 1302 15 AK102641
15 209 39.2 46422 15 OSJN00249
16 209 39.2 90825 15 AC157500
17 209 39.2 110000 15 AP008210_221
18 209 39.2 118959 15 OSJN00235
19 197.5 37.1 110000 15 AP008218_057
20 197.5 37.1 159056 15 CNS080C7
21 179.5 33.7 117439 15 AC122163
22 179.5 33.7 154378 15 AC126006
23 177 33.2 154378 15 AK070163
24 177 33.2 110000 15 AP008207_386
25 177 33.2 191022 15 AP003431
26 170.5 32.0 629 15 DQ099057
27 167.5 31.4 77319 15 AP004474
28 167.5 31.4 91654 15 AP006393
29 164.5 30.9 352 10 BV144464
30 163.5 30.7 107407 15 AC134933
31 163.5 30.7 110000 15 AP008211_203
32 157 29.5 559 15 AK070327
33 157 29.5 1044 15 AK061571
34 157 29.5 110000 15 AP008214_007
35 157 29.5 185545 15 AP005657
36 153.5 28.8 330 10 BV146385
37 153.5 28.8 343 10 BV146386
38 153.5 28.8 345 10 BV146389
39 151 28.3 111256 15 AC125478
40 149.5 28.0 342 10 BV146383
41 149.5 28.0 342 10 BV146384
42 149.5 28.0 342 10 BV146388
43 147.5 27.7 342 10 BV146393
44 145 27.2 32163 14 AC160455
45 144.5 27.1 326 10 BV146392

ALIGNMENTS

RESULT 1
BT011724
LOCUS Arabidopsis thaliana At3g59900 gene, complete cds.
DEFINITION
ACCESSION BT011724
VERSION BT011724.1 GI:44917524
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
1 (bases 1 to 393)
REFERENCE
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Carninci, P., Hayashizaki, Y.,
Ishida, J., Kamiya, A., Kawai, J., Narisaka, M., Sakurai, T., Satou, M.,
Seki, M., Shinozaki, K., and Ecker, J.R.
TITLE Arabidopsis ORF clones
JOURNAL
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Shinn, P., Carninci, P., Hayashizaki, Y.,
Ishida, J., Kamiya, A., Kawai, J., Narisaka, M., Sakurai, T., Satou, M.,
Seki, M., Shinozaki, K., and Ecker, J.R.
TITLE Direct Submission
JOURNAL
AUTHORS Submitted (04-MAR-2004) Salk Institute Genomic Analysis Laboratory
(SIGAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PSCC (SSP) Consortium members constructed and sequenced the pJN1 (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Shim, P., Bowser, L., Chan, M. M., Chang, C. M., Dale, J. M., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

CDS

1..393
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="U65994"
/ecotype="Columbia"
1..393
/note="unknown protein"
/product="At3g5990"
/protein_id="AA049087.1"
/db_xref="GI:44917525"
/translation="MIRISNLOKDIINIDSYNNRVMDVGNRRKNSFRSPSKS
KQELRSFSAOKMIPANYPFSLSLFLVGLTASLLILPLVPLPPFPLLVPI
GIWLVVLAFMPSHSHNANTDVTGCMF"

ORIGIN

Alignment Scores:

Pred. No.: 1,11e-45 Length: 393
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-715-129-2 (1-106) x BT011724 (1-393)

QY 1 MetAspValGlyArgAsnAsnArgLyAsnMetSerPheArgSerSerProGluLySer 20
Db 73 ATGAGCGTCGGAAGAAACAACCGGAAACATGAGCTTTGGAAGTTGCCGGAAGAAAGC 132
QY 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLyArgMetMetIleProAlaAsnTyr 40
Db 133 AAGCAAGAGTTACGCGGAGTTCTCGCGCAAAAGATGATGATCCGCGCAATTAT 192
QY 41 PheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeuPro 60
Db 193 TTCAGTTAGAGCTCTGTTCCATGCTGTTGCTTAACGCGCATCTCTGTTAATACCTCCG 252
QY 61 LeuValLeuProProLeuProProProProPheMetLeuLeuValProIleGlyIle 80
Db 253 TTAGTTTGGCGCGGTTACCTCCGCTTCGTTATGCTGCTATTTGGTCCATTTGGGATT 312
QY 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
Db 313 ATGGTTTACTCGCGCTTCTTGCTTCAATGCTTCTTCTCAATCTTAATGCTAATACAGAT 372
QY 101 ValThrCysAsnPheMet 106
Db 373 GTAACCTGCAATTCATG 390

RESULT 2
AY305869
LOCUS

AY305869 732 bp mRNA linear PLN 03-SEP-2003

DEFINITION

Arabidopsis thaliana auxin-inducible protein (ARGOS) mRNA, complete cds.

ACCESSION
AY305869
VERSION
AY305869.1
KEYWORDS
GI:34329825

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE

1 (bases 1 to 732)

AUTHORS

Hu, Y., Xie, Q. and Chua, N.-H.

TITL

The Arabidopsis Auxin-Inducible Gene ARGOS Controls Lateral Organ

JOURNAL

Plant Cell 15 (9), 1951-1961 (2003)

PUBMED

12953103

REFERENCE

2 (bases 1 to 732)

AUTHORS

Hu, Y., Xie, Q. and Chua, N.-H.

TITL

Submitted (24-MAY-2003) Laboratory of Molecular Cell Biology,

JOURNAL

Temasek Life Sciences Laboratory, 1 Research Link, National

FEATURES

University of Singapore, Singapore 117604, Singapore

source

Location/Qualifiers
1..732
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/map="At3g59900"
/ecotype="Columbia"
1..732
/gene="ARGOS"
388..708
/gene="ARGOS"
/codon_start=1
/product="auxin-inducible protein"
/protein_id="AA064004.1"
/db_xref="GI:34329826"
/translation="MDVGNRRKNSFRSPSKSKQELRSFSAOKMIPANYPFSL
SEFLVGLTASLLILPLVPLPPFPLLVPIGIMVLVLAFMPSHSHNANTDVT
GCMF"

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.: 2,12e-45 Length: 732
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-715-129-2 (1-106) x AY305869 (1-732)

QY 1 MetAspValGlyArgAsnAsnArgLyAsnMetSerPheArgSerSerProGluLySer 20
Db 388 ATGAGCGTCGGAAGAAACAACCGGAAACATGAGCTTTGGAAGTTGCCGGAAGAAAGC 447
QY 21 LysGlnGluLeuArgArgSerPheSerAlaGlnLyArgMetMetIleProAlaAsnTyr 40
Db 448 AAGCAAGAGTTACGCGGAGTTCTCGCGCAAAAGATGATGATCCGCGCAATTAT 507
QY 41 PheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeuPro 60
Db 508 TTCAGTTAGAGCTCTGTTCCATGCTGTTGCTTAACGCGCATCTCTGTTAATACCTCCG 567
QY 61 LeuValLeuProProLeuProProProProPheMetLeuLeuValProIleGlyIle 80
Db 568 TTAGTTTGGCGCGGTTACCTCCGCTTCGTTATGCTGCTATTTGGTCCATTTGGGATT 627
QY 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
Db 628 ATGGTTTACTCGCGCTTCTTGCTTCAATGCTTCTTCTCAATCTTAATGCTAATACAGAT 687

Qy 101 ValThrcCyasnpHmet 106
 Db 688 GTAACCTGCATTTTCATG 705
 RESULT 3
 AKI16640
 LOCUS
 DEFINITION
 AKI16640 747 bp mRNA linear PLN 09-SEP-2004
 clone: RAFL25-17-J07.
 ACCESSION
 AKI16640.1 GI:51971476
 VERSION
 FLI CDNA.
 KEYWORDS
 Arabidopsis thaliana (thale cress)
 SOURCE
 Arabidopsis thaliana
 ORGANISM
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE
 AUTHORS
 Totsuki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K., Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K., Kamiya, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
 Unpublished
 2 (bases 1 to 747)
 Totsuki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K., Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K., Kamiya, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Direct Submission
 Submitted (06-SEP-2004) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: msekik@gscc.riken.jp, URL: http://xlarge.gsc.riken.jp/, Tel: 81-45-503-9625, Fax: 81-45-503-9586)
 COMMENT
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720, Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified lambdaBda FLC-1-B vector (Carninci et al. (2001) Genome 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector.
 Please visit our web site (http://xlarge.gsc.riken.jp/) for further details.
 FEATURES
 source
 1. .747
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL25-17-J07"
 /ecotype="Columbia"
 /note="common name: thale cress"
 1. .747
 /gene="At3g59900"
 314. .706
 /gene="At3g59900"
 /codon_start=1
 /product="putative protein"
 /protein_id="BAD44403.1"
 /db_xref="GI:51971477"
 /translation="MIRISLNLQDIINIDSYNNRVMVDYGRNNRKNMSFRSSPEKS KQIRLSPSAQKRWMIIPANYSLSLFLVGLTASLILPLVPLPPPPMLLVPI GIVLVVLAIFMPSHSNANTDVCNMF"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.17e-45 Length: 747
 Score: 533.00 Matches: 106
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0

DB: 15 Gaps: 0
 US-10-715-129-2 (1-106) x AKI16640 (1-747)
 Qy 1 MetNapValGlyArgAaenAaArgLyAaenMetSerPheArgSerProGlyLysSer 20
 Db 386 ATGACGCTCGGAAGAAACCAACCGGAAAAACATGAGCTTTCGAGATTCGCCGGAAGAAC 445
 Qy 21 LygGlnGluLeuAaGArgSerPheSerAlaGlyArgMetMetIleProAlaAsnTyr 40
 Db 446 AAGCAAGATTACCGCCGAGTTTCTCCGCGCAAGAAAGATGATGATCCCGCAATTTAT 505
 Qy 41 PheSerLeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuIleLeuPro 60
 Db 506 TTCAGTTAGAGTCTCTGTTCTTATGTTGTTACCGCATCTCTGTTAACTTCCG 565
 Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuLeuValProIleGlyTle 80
 Db 566 TTAGTTTGGCGCGCTTACCTCGCGCTCCGTTTATGCTGATTTGTTCCATGGAGTT 625
 Qy 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAlaAsnThrAsp 100
 Db 626 ATGCTTTTACTCGTCGCTTCTTGCCCTTCATGCTTCTCATTTCAATGCTAAATACAGAT 685
 Qy 101 ValThrcCyasnpHmet 106
 Db 686 GTAACCTGCATTTTCATG 705
 RESULT 4
 ATP24G16
 LOCUS
 DEFINITION
 Arabidopsis thaliana DNA chromosome 3, BAC clone F24G16.
 ACCESSION
 AL138647
 VERSION
 AL138647.1 GI:6899904
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE
 AUTHORS
 D'Angelo, M., Vezzi, A., Modesto, D., Pigazzi, M., Valle, G., Mewes, H.W., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.
 Unpublished
 2 (bases 1 to 100906)
 EU Arabidopsis sequencing project.
 DIRECT SUBMISSION
 Submitted (18-FEB-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project Coordinator: Marcel Salanoubat and Francis Quetier, Groupement d'interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cremieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr
 COMMENT
 Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
 FEATURES
 source
 1. .100906
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /ecotype="Columbia"
 1. .1731
 /note="overlap to BAC F16L24, please refer EMBL Acc. no AL138659 for analysis and annotation"
 2. .1888
 /gene="F24G16.10"
 /gene="F24G16.10"
 /gene="F24G16.10"
 complement (2. .1888)
 /gene="F24G16.10"
 complement (2. .1888)
 /gene="F24G16.10"

exon complement (7549, .7658)
/gene="F24G16.30"
/number=9
intron complement (7659, .7818)
/gene="F24G16.30"
/number=9
exon complement (7819, .8175)
/gene="F24G16.30"
/number=10
gene 1211, .18504
/gene="F24G16.40"

Alignment Scores:

Pred. No.: 3,49e-43 Length: 100906
Score: 533.00 Matches: 106
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-715-129-2 (1-106) x ATF24G16 (1-100906)

Qy 1 MetAspValGlyArgAsnAsnArglyAsnMetSerPheArgSerSerProGlnlySer 20
Db 62943 ATGACCGTGGAGAAACACCGGAAAACATGACCTTCCAGTTCCCGGAGAAAAGC 63002
Qy 21 LysGlnGlnLeuArgArgSerPheSerAlaGlnlySerMetMetIleProAlaSer 40
Db 63003 AAGCAAGATTACGCGGAGTTTCTCGGCGCAAAAAGATGATGATCCCGGCAATTAT 63062
Qy 41 PheSerLeuGlnSerIlePheLeuLeuValGlyLeuThrAlaSerLeuLeuIleLeuPro 60
Db 63063 TTGAGTTTAGAGCTCTGTTCTTATGGTGTCTTAAGGCAATCTCTGTATATCTTCG 63122
Qy 61 LeuValLeuProProLeuProProProPheMetLeuLeuValProIleGlyIle 80
Db 63123 TTAGTTTGGCCGCGCTTACCTCCGCTTATGCTGCTATGTGTTCCCATTTGGAGTT 63182
Qy 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaSerThrAsp 100
Db 63183 ATGGTTTATCGTCGCTTCTTGCCTTCAATGCTTCTTCAATTCATTAATGATATGCAAT 63242
Qy 101 ValThrCysAsnPheMet 106
Db 63243 GTAACCTGCATTTTCATG 63260

RESULT 5
AX506572 LOCUS AX506572 336 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1267 from Patent WO0216655.
ACCESSION AX506572
VERSION AX506572.1 GI:23387809
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 Harper, J.F., Krebs, J., Wang, X. and Zhu, T.,
AUTHORS Stress-regulated genes of plants, transgenic plants containing
TITLE same, and methods of use
JOURNAL Patent: WO 0216655-A 1267 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source location/Qualifiers
1..336
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

Alignment Scores:

Pred. No.: 1.3e-24 Length: 336
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 6 Gaps: 3

US-10-715-129-2 (1-106) x AX506572 (1-336)

Qy 1 MetAspValGlyArgAsnAsnArglyAsnMetSerPheArgSerSerProGlnlySer 16
Db 1 ATGACCGTGGAGAAACACCGGAAAACATGACCTTCCAGTTCCCGGAGAAAAGC 60
Qy 17 ProGlnlySerIleGlnGlnLeuArgArgSerPheSerAlaGlnlySerArgMet 34
Db 61 CCGATTATGGGAGCAAGAAATTTTTCGACATTTGCTGTCGCAACAGTCCAGAGAGG 120
Qy 35 MetIleProAlaSerIlePheSerLeuGlnSerIlePheLeuLeuValGlyLeuThrAla 54
Db 121 CTATATACAGCGAGTTACTTCACTTCAATCAATGCTTGCTTGTGCTCAGCA 180
Qy 55 SerLeuLeuIleLeuProLeuValIleProProLeuProProProPheMetLeuLeu 74
Db 181 TCTCTTGAATCTTACCGTTGATCTTCCACATTTGCTCTCTCTTATGCTGCTT 240
Qy 75 LeuValProIleGlyIleMetValLeuLeuValValLeuAlaPheMetProSerSerHis 94
Db 241 TTGATTCCTATGGGATTAATGTTTGTGTTATGTTCTTCAATGCTTCTTCAAT 300
Qy 95 Ser 95
Db 301 TCC 303

RESULT 6
AX651239 LOCUS AX651239 336 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 23 from Patent WO0300898.
ACCESSION AX651239
VERSION AX651239.1 GI:29154057
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
AUTHORS Karagiannis, F., Qian, S., Tao, Y., Whitman, S., Xie, Z., Zhu, T. and Zou, G.,
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 0300898-A 23 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source location/Qualifiers
1..336
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

Alignment Scores:

Pred. No.: 1.3e-24 Length: 336
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 6 Gaps: 3

US-10-715-129-2 (1-106) x AX651239 (1-336)

Qy 1 MetAspValGlyArgAsnAsnArglyAsnMetSerPheArgSerSerProGlnlySer 16
Db 1 ATGACCGTGGAGAAACACCGGAAAACATGACCTTCCAGTTCCCGGAGAAAAGC 60
Qy 17 ProGlnlySerIleGlnGlnLeuArgArgSerPheSerAlaGlnlySerArgMet 34

```

Db      61 CCGATTATGGGGAAGCAAGATTGTTTCGACATGTCGTGCAGAAACATGCCAAGGAGG 120
Qy      35 MetIleProIlaaenTyRphSerLeuGluSerLeuPheLeuValGlyLeuThrIa 54
Db      121 CTAATATCAGGAGATTAAGTTCAGTTAGAAATCAATGGTTGGCTGTGGCTCACAACA 180
Qy      55 SerLeuLeuIleuProLeuValIleuProProLeuProProProPheMetLeuIleu 74
Db      181 TCTCTTGAATCTTACGCTGATCTTCACCAATGCTCTCTCTCTTATATCTCTT 240
Qy      75 LeuValProIleGlyIleMetValIleuValIleuValIleuValaPheMetProSerSerHis 94
Db      241 TTGATTCCTATGGGATATGATGTTTGCTATGCTTCTGCTTCATGCTCTTCTAT 300
Qy      95 Ser 95
Db      301 TCC 303

RESULT 7
BT000930
LOCUS   Arabidopsis thaliana clone C105228 unknown protein (At2g44080)
DEFINITION
ACCESSION   BT000930
VERSION     BT000930.1 GI:24030324
KEYWORDS   FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana

REFERENCE
AUTHORS     Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
            Nguyen,M., Palm,C.J., Shin,P., Southwick,A., Tripp,M.G., Wu,T.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Arabidopsis Open Reading Frame (ORF) Clones
            Unpublished
            2 (bases 1 to 439)

TITLE       JOURNAL
REFERENCE   2 (bases 1 to 439)
AUTHORS     Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
            Nguyen,M., Palm,C.J., Shin,P., Southwick,A., Tripp,M.G., Wu,T.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Direct Submission
            Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
            Annotation based on January 2002 version of the Arabidopsis genome
            submitted to Genbank.

FEATURES
SOURCE
location/Qualifiers
1..439
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="C105228"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1..439
/gene="At2g44080"
1..408
/gene="At2g44080"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA041330.1"
/db_xref="GI:24030325"
/translation="MIREPSLDNDIINIOEHSYNNMVRGDRHAKTSFSGSAPA
PIKQDELFTLSQNSPRRLISAYSLSMVVLVLTASLLILPLPLPPPPPM
LTLIPGIVLWLVLAWLPSSNSKRVSSSTFM"
409..439
3' UTR

```

```

ORIGIN                                           /gene="At2g44080"

Alignment Scores:
Pred. No.: 1,72e-24 Length: 439
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 15 Gaps: 3

US-10-715-129-2 (1-106) x BT000930 (1-439)
Qy      1 MetAspVal---GlyArgaanaArgLysaAmetSerPheArgSerSer----- 16
Db      73 ATGACCTGAGAGAGAGATCAATACCGAANAACACGAGTTTGTGGTTCAAGCTCAGCT 132
Qy      17 ProGluysSerLysGlnGluLeuArgArgSerPheSerLagInlys-----ArgMet 34
Db      133 CCGATTATGGGGAAGCAAGATTGTTTCGACATGTCGTGCAGAAACATGCCAAGGAGG 192
Qy      35 MetIleProIlaaenTyRphSerLeuGluSerLeuPheLeuValGlyLeuThrIa 54
Db      193 CTAATATCAGGAGATTAAGTTCAGTTAGAAATCAATGGTTGCTGTGGCTCACAACA 252
Qy      55 SerLeuLeuIleuProLeuValIleuProProLeuProProProPheMetLeuIleu 74
Db      253 TCTCTTGAATCTTACGCTGATCTTCACCAATGCTCTCTCTCTTATATCTCTT 312
Qy      75 LeuValProIleGlyIleMetValIleuValIleuValIleuValaPheMetProSerSerHis 94
Db      313 TTGATTCCTATGGGATATGATGTTTGCTATGCTTCTGCTTCATGCTCTTCTAT 372
Qy      95 Ser 95
Db      373 TCC 375

RESULT 8
AY080817
LOCUS   Arabidopsis thaliana unknown protein (At2g44080) mRNA, complete
DEFINITION
ACCESSION   AY080817
VERSION     AY080817.1 GI:19423989
KEYWORDS   FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana

REFERENCE
AUTHORS     Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Setou,M., Seki,M., Shin,P., Southwick,A., Shinzaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Arabidopsis Full Length cDNA Clones
            Unpublished
            2 (bases 1 to 904)

TITLE       JOURNAL
REFERENCE   2 (bases 1 to 904)
AUTHORS     Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,X.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Setou,M., Seki,M., Shin,P., Southwick,A., Shinzaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
            Direct Submission
            Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
            RIKEN Genomic Sciences Center (GSC) members carried out the
            COMMENT

```


gene /rptc_family="AT_rich"
7192..8658
/gene="At2g43890"
/note="synonym: F6B13.2"
join(<7192..7632,7743..8032,8122..8350,8440..>8658)
/gene="At2g43890"
join(7192..7632,7743..8032,8122..8350,8440..8658)
/gene="At2g43890"
/product="putative polygalacturonase"
/protein_id="AAC23398.1"
/db_xref="GI:3212847"
/translation="MDNNKLVAVLLMFSSFLMLKSTASTANVVSFGAKPDGRTDS
TKAFGLMOPACRASAAYVTVPGRSLPVPFGRCPRIFFOYGLTVASDVRS
LNGSWYILFVAKNRIISIGTLDARASFMARCKSKSGSPRTFMADVVVS
GLTINSOTHLVINSQNVIVRYKVLVADPSDPTDGLVOSAGTVLDPFHGD
DCISIGDTNLTMSKLNCGPGHGISIGSLGRANAGVNTLINSVPSGSDNGRI
KTAAROSTGFNRNVLFONLTKNVONFIYDNYCPBNQCPKQSGVKSIVVYENI
QGRSTQOALTFCDSRSNPCOAILRHDIKLTFGRSATSTCKNIKGVKGVNMPQCL
"
CDS complement(9101..14090)
/gene="At2g43900"
/note="synonym: F6B13.3"
complement(join(<9101..9952,10049..10409,10500..10623,
10717..10984,11093..11133,11219..11384,11587..11671,
11768..11895,12004..12811,12910..13162,13259..>14090))
/gene="At2g43900"
complement(join(9101..9952,10049..10409,10500..10623,
10717..10984,11093..11133,11219..11384,11587..11671,
11768..11895,12004..12811,12910..13162,13259..14090))
/gene="At2g43900"
/codon_start=1
/product="putative inositol polyphosphate 5'-phosphatase"
/protein_id="AAC23399.1"
/db_xref="GI:3212848"
/translation="MDIINNHRDENDDEEALSAMSVPPRKTHSYHQIARQ
KGRHORSHLDIIPKITEIVSGGIGSDSDDEFPYATTTSSSFPFGDPTGSD
DYLOPEIGEDFQPLPEFVSGGVGAFKPTSPILSARPCLELRFKPTQWR
FLINACTETOLMAGOSGVRFMNFDAFPGGLSGRVORGEDAPFESASTPT
TCMLVNGNRLVMSGHDKIRIKMDYVLDGDDSPFKGSLMOAKHQPVNSVINS
YGVLMSCSEGVYKIKTWBSMEKSLRLTEKEMALLVRSGLIDRAQVTVGTCT
SSSEVKCLADNVRSKVMAQLOTFSLMDRTKLELVNFBEGQTERVMPAGQDP
LAADENKAKIATGTSKKEKPGFLORSNANAGADAVRVAIVGGAGYEDAKTEMA
LADGNLMTGCTGTLIQMDGNRLQDFRHHOCAYVCFCTPERIYIIGVSHILOI
DLEGNLAVGNHNNAYIKMAADGYFSLATGSGRGMPVISRGLDGIIRSELK
ERTYAOITDSVRLITGSMNVGOGKASHDALMSYGSVADYGLVGLQYREMGAKLA
MSAKESVGNESSTIGQWIDTIGKTLDEKAVPERKSGQLAGLISLWVKRLKTH
VGIIDVAVPCGFRALGNKGVGRLRIVERDRIMCFINCHLAHLEAVNRNADPHI
YKMSFTSRSSNANAPAGVSTGSHTKSANNAVNTETKODLEADVVPFGDPT
RLFGIYSDERDVFVSGRFDMLREKQDLRAEMKAGRVFGMRBAITFPYTFEERH
PGAGYDSGEKKRIIPACDRVIFPDRTSPESGSLDPCPVASIMLYDAMDYESH
KVPKRCFKVKTIEHVDSVRROEFGRIKTMEXKRALINDLRYVETIVSNGSLVQNO
DTEVLRITTKCKENAVFRILCBQSTVRDEDTLELPLSGSGFPMELVMPAAGT
KIPSSVEVSVHHEFHTLEFVDGIPONNWCEBTRDEALLVNVOGSGSTEVCHRY
HVHCFEAKNLRIDSNPSNKSQSLKKNBEDSNKSKSKSDGNSNKSCKSGSDNS
SKSKSKSDGNSKSKSDGNSKSKSKSDGNSKSKSDGNSKSKSDGNSKSKSDGNS
CSKSKSDGNTNSKSKSDGNSKSKSKSHKNDGDSKSKHKNKGSSSKSHKSKSDG
DSSSKSHKSKSGSLSLTRITRTEILLQAHTRSRVGRKTVTRQL"
repeat_region 14103..14213
/rptc_family="GAGAA)n"
complement(14748..17292)
/gene="At2g43910"
/note="synonym: F6B13.4; supported by full length cDNA:
Ceres.24370"
complement(join(<14748..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..>17292))
/gene="At2g43910"
complement(join(14925..14963,15285..15363,15476..15522,
15617..15716,15874..15974,16059..16145,16677..16807,
17119..17218))
/gene="At2g43910"
/codon_start=1

repeat_region /product="putative methyl chloride transferase"
/protein_id="AAC23400.1"
/db_xref="GI:3212849"
/translation="MAEEOQNSQSGNVIPREPARATKHTVEGCMKEREI
TPMDQARATPLVHLVDITSLPLGRALVPGCGGHDVMAASPERVVGIDISEBAL
KANETYGSSPKAEYFSEFVEDVETMRPDLIFDYVFCAIEPMRPAMAKMTEL
LKPDEGLITLMPITDHVGGPPYKVDVSTFEELVPIGFKAIVSEBNPAIPYRKKE
KLRWKKIN"
complement(15108..15141)
/rptc_family="AT_rich"
repeat_region complement(16570..16608)
/rptc_family="AT_rich"
repeat_region complement(17642..17687)
/rptc_family="AT_rich"
repeat_region 17774..17794
/rptc_family="AT_rich"
repeat_region 17926..17970
/rptc_family="AT_rich"
repeat_region 18152..18196
/rptc_family="AT_rich"
complement(19342..21515)
/gene="At2g43920"
/note="synonym: F6B13.5"
complement(join(<19342..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..>21515))
/gene="At2g43920"
complement(join(19586..19624,19773..19851,19988..20034,
20130..20229,20308..20408,20521..20607,20828..20958,
21416..21515))
/gene="At2g43920"
/codon_start=1
/product="unknown protein"
/protein_id="AAC23401.1"
/db_xref="GI:3212850"
/translation="MAEEOQNSQSGNIPPTPEBAATFQPVVAEGWCKMEDGV
TPMDQARATPLVHLVDITSLPLGRALVPGCGGHDVMAASPERVVGIDISEBAL
KANETYGSSPKAEYFSEFVEDVETMRPDLIFDYVFCAIEPMRPAMAKMTEL
LKPDEGLITLMPITDHVGGAPYKVALSYEDLVVGFVAIVSEBNPDSIPYRKKE
KLARKKIN"
23418..25849
/gene="At2g43930"
/note="synonym: F6B13.6; contains a protein kinase domain
profile (PDOC00100)"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
join(23418..23599,24062..24130,24488..24592,24751..24812,
24898..25008,25727..25849)
/gene="At2g43930"
/codon_start=1
/product="putative protein kinase"
/protein_id="AAC23427.1"
/db_xref="GI:3212876"
/translation="MDHHRRLATIKLIRNSSTLLFPKIVPILHRSGDGRALPYT
PFPSICRALAGHNTSRSEPHVILLOCKRKSVCPSSEFELCLFGYIDEDERKHF
GHKGVVAAREEVOLIKNSHPNIVYIDITVRSDSLNTMMQVPGG8168LLAKTGS
FPEPKHNRVLDNAITSHYSYKNIHQOSFKLEIDPRKKGS"
complement(27269..30414)
/gene="At2g43940"
/note="synonym: F6B13.7"
complement(join(<27269..27361,27641..27771,27853..27984,
28336..28397,28479..28537,28635..28784,28865..28981,
29149..29343,30043..>30414))
/gene="At2g43940"

Alignment Scores:
Pred. No.: 5,23e-22 Length: 109741
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: 15 Gaps: 3

[illegible]

JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skkuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.

COMMENT

URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Negata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yezaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Onoda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shibasaki,T. and
Yamamoto,M.
PAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumagi,T., Lu,M., Maeda,H., Miura,J.,
Mizuno,K., Narikawa,R., Nikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsumoda,Y., Ueda,M., Xie,Q., Yokoiizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hata,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hitaka,T., Horii,F., Iida,J., Imanure,K., Imotani,K., Ishi,Y.,
Itou,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawat,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondou,S., Konno,H., Kouda,M.,
Koys,S., Kurahara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Nunaseki,R., Ohno,M., Otsuo,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Saeki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takahata-Akihita,S., Tanaka,Y., Tomaru,A., Toya,T., waki,K.,
Yaenunishi,A. and Hayashizaki,Y.
Location/Qualifiers
1..1137
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-021-P08"

ORIGIN

Alignment Scores:

| Pred. No.: | 1.Sle-13 | Length: | 1137 |
|------------------------|----------|---------------|------|
| Score: | 224.50 | Matches: | 51 |
| Percent Similarity: | 65.7% | Conservative: | 16 |
| Best Local Similarity: | 50.0% | Mismatches: | 24 |
| Query Match: | | Indels: | 11 |
| Ds: | 15 | Gaps: | 3 |

US-10-715-129-2 (1-106) x AK060546 (1-1137)

Oy 3 ValGtGAgaAaaAhArgLyahsmetSerPheaGserSerProdlulySerlygIn 22
Db 668 GTAGGTGCTTAGGCAGAGAAGAAAC---CCTCTGATATCCCAAGCAAAGCTTCAAGAG 744
Oy 23 GlueUAtGAgsrPhe-----SerAlngLntYAraMetMetlIEProAlaAn 39
Db 745 CATCTGCAGCATCACATCAACCCTAAATGCTGCCAGAAAGAAACCGCCGACATCGAAT 804
Oy 40 TyPheSerLeuGlusierLeuPheLeuValglyleuthralaserLeuileu 59
Db 805 TACTTCAGATGAGAGCGCTTCGTGGCTCGTCTTCCTCACATGCTATGCTCATACTT 864
Oy 60 ProLeuValleuProProleuProPropoPhemeLeuleuValProlegly 79
Db 865 CCATTGGTGCTCCCCCAATGCTCCGCGCCATCGTCTGCTGCTGCCAGTCTGC 924
Oy 80 lIeetValleuLeuValleuAlaPheMetProSerSerHisserAnaIaenth 99
Db 925 CTGCTCATCTGCTGATTTGCTGGCTTCATGCCA-----ACG 963

100 AspvAl 101

Db 964 GATGTG 969

RESULT 12
AP008217_078/c
WPCOMMENT

Sequence split into 264 fragments LOCUS AP008217 Accession AP008217

| Fragment Name | Begin | End |
|---------------|---------|---------|
| AP008217_000 | 1 | 110000 |
| AP008217_001 | 100001 | 210000 |
| AP008217_002 | 200001 | 310000 |
| AP008217_003 | 300001 | 410000 |
| AP008217_004 | 400001 | 510000 |
| AP008217_005 | 500001 | 610000 |
| AP008217_006 | 600001 | 710000 |
| AP008217_007 | 700001 | 810000 |
| AP008217_008 | 800001 | 910000 |
| AP008217_009 | 900001 | 1010000 |
| AP008217_010 | 1000001 | 1110000 |
| AP008217_011 | 1100001 | 1210000 |
| AP008217_012 | 1200001 | 1310000 |
| AP008217_013 | 1300001 | 1410000 |
| AP008217_014 | 1400001 | 1510000 |
| AP008217_015 | 1500001 | 1610000 |
| AP008217_016 | 1600001 | 1710000 |
| AP008217_017 | 1700001 | 1810000 |
| AP008217_018 | 1800001 | 1910000 |
| AP008217_019 | 1900001 | 2010000 |
| AP008217_020 | 2000001 | 2110000 |
| AP008217_021 | 2100001 | 2210000 |
| AP008217_022 | 2200001 | 2310000 |
| AP008217_023 | 2300001 | 2410000 |
| AP008217_024 | 2400001 | 2510000 |
| AP008217_025 | 2500001 | 2610000 |
| AP008217_026 | 2600001 | 2710000 |
| AP008217_027 | 2700001 | 2810000 |
| AP008217_028 | 2800001 | 2910000 |
| AP008217_029 | 2900001 | 3010000 |
| AP008217_030 | 3000001 | 3110000 |
| AP008217_031 | 3100001 | 3210000 |
| AP008217_032 | 3200001 | 3310000 |
| AP008217_033 | 3300001 | 3410000 |
| AP008217_034 | 3400001 | 3510000 |
| AP008217_035 | 3500001 | 3610000 |
| AP008217_036 | 3600001 | 3710000 |
| AP008217_037 | 3700001 | 3810000 |
| AP008217_038 | 3800001 | 3910000 |
| AP008217_039 | 3900001 | 4010000 |
| AP008217_040 | 4000001 | 4110000 |
| AP008217_041 | 4100001 | 4210000 |
| AP008217_042 | 4200001 | 4310000 |
| AP008217_043 | 4300001 | 4410000 |
| AP008217_044 | 4400001 | 4510000 |
| AP008217_045 | 4500001 | 4610000 |
| AP008217_046 | 4600001 | 4710000 |
| AP008217_047 | 4700001 | 4810000 |
| AP008217_048 | 4800001 | 4910000 |
| AP008217_049 | 4900001 | 5010000 |
| AP008217_050 | 5000001 | 5110000 |
| AP008217_051 | 5100001 | 5210000 |
| AP008217_052 | 5200001 | 5310000 |
| AP008217_053 | 5300001 | 5410000 |
| AP008217_054 | 5400001 | 5510000 |
| AP008217_055 | 5500001 | 5610000 |
| AP008217_056 | 5600001 | 5710000 |
| AP008217_057 | 5700001 | 5810000 |
| AP008217_058 | 5800001 | 5910000 |
| AP008217_059 | 5900001 | 6010000 |
| AP008217_060 | 6000001 | 6110000 |
| AP008217_061 | 6100001 | 6210000 |
| AP008217_062 | 6200001 | 6310000 |
| AP008217_063 | 6300001 | 6410000 |
| AP008217_064 | 6400001 | 6510000 |

| | | |
|--------------|----------|----------|
| AP008217_065 | 6500001 | 6610000 |
| AP008217_066 | 6600001 | 6710000 |
| AP008217_067 | 6700001 | 6810000 |
| AP008217_068 | 6800001 | 6910000 |
| AP008217_069 | 6900001 | 7010000 |
| AP008217_070 | 7000001 | 7110000 |
| AP008217_071 | 7100001 | 7210000 |
| AP008217_072 | 7200001 | 7310000 |
| AP008217_073 | 7300001 | 7410000 |
| AP008217_074 | 7400001 | 7510000 |
| AP008217_075 | 7500001 | 7610000 |
| AP008217_076 | 7600001 | 7710000 |
| AP008217_077 | 7700001 | 7810000 |
| AP008217_078 | 7800001 | 7910000 |
| AP008217_079 | 7900001 | 8010000 |
| AP008217_080 | 8000001 | 8110000 |
| AP008217_081 | 8100001 | 8210000 |
| AP008217_082 | 8200001 | 8310000 |
| AP008217_083 | 8300001 | 8410000 |
| AP008217_084 | 8400001 | 8510000 |
| AP008217_085 | 8500001 | 8610000 |
| AP008217_086 | 8600001 | 8710000 |
| AP008217_087 | 8700001 | 8810000 |
| AP008217_088 | 8800001 | 8910000 |
| AP008217_089 | 8900001 | 9010000 |
| AP008217_090 | 9000001 | 9110000 |
| AP008217_091 | 9100001 | 9210000 |
| AP008217_092 | 9200001 | 9310000 |
| AP008217_093 | 9300001 | 9410000 |
| AP008217_094 | 9400001 | 9510000 |
| AP008217_095 | 9500001 | 9610000 |
| AP008217_096 | 9600001 | 9710000 |
| AP008217_097 | 9700001 | 9810000 |
| AP008217_098 | 9800001 | 9910000 |
| AP008217_099 | 9900001 | 10010000 |
| AP008217_100 | 10000001 | 10110000 |
| AP008217_101 | 10100001 | 10210000 |
| AP008217_102 | 10200001 | 10310000 |
| AP008217_103 | 10300001 | 10410000 |
| AP008217_104 | 10400001 | 10510000 |
| AP008217_105 | 10500001 | 10610000 |
| AP008217_106 | 10600001 | 10710000 |
| AP008217_107 | 10700001 | 10810000 |
| AP008217_108 | 10800001 | 10910000 |
| AP008217_109 | 10900001 | 11010000 |
| AP008217_110 | 11000001 | 11110000 |
| AP008217_111 | 11100001 | 11210000 |
| AP008217_112 | 11200001 | 11310000 |
| AP008217_113 | 11300001 | 11410000 |
| AP008217_114 | 11400001 | 11510000 |
| AP008217_115 | 11500001 | 11610000 |
| AP008217_116 | 11600001 | 11710000 |
| AP008217_117 | 11700001 | 11810000 |
| AP008217_118 | 11800001 | 11910000 |
| AP008217_119 | 11900001 | 12010000 |
| AP008217_120 | 12000001 | 12110000 |
| AP008217_121 | 12100001 | 12210000 |
| AP008217_122 | 12200001 | 12310000 |
| AP008217_123 | 12300001 | 12410000 |
| AP008217_124 | 12400001 | 12510000 |
| AP008217_125 | 12500001 | 12610000 |
| AP008217_126 | 12600001 | 12710000 |
| AP008217_127 | 12700001 | 12810000 |
| AP008217_128 | 12800001 | 12910000 |
| AP008217_129 | 12900001 | 13010000 |
| AP008217_130 | 13000001 | 13110000 |
| AP008217_131 | 13100001 | 13210000 |
| AP008217_132 | 13200001 | 13310000 |
| AP008217_133 | 13300001 | 13410000 |
| AP008217_134 | 13400001 | 13510000 |
| AP008217_135 | 13500001 | 13610000 |
| AP008217_136 | 13600001 | 13710000 |
| AP008217_137 | 13700001 | 13810000 |

AP008217_138 13800001 13910000
 AP008217_139 13900001 14010000
 AP008217_140 14000001 14100000
 AP008217_141 14100001 14200000
 AP008217_142 14200001 14300000
 AP008217_143 14300001 14400000
 AP008217_144 14400001 14500000
 AP008217_145 14500001 14600000
 AP008217_146 14600001 14700000
 AP008217_147 14700001 14800000
 AP008217_148 14800001 14900000
 AP008217_149 14900001 15000000
 AP008217_150 15000001 15100000
 AP008217_151 15100001 15200000
 AP008217_152 15200001 15300000
 AP008217_153 15300001 15400000
 AP008217_154 15400001 15500000
 AP008217_155 15500001 15600000
 AP008217_156 15600001 15700000
 AP008217_157 15700001 15800000
 AP008217_158 15800001 15900000
 AP008217_159 15900001 16000000
 AP008217_160 16000001 16100000
 AP008217_161 16100001 16200000
 AP008217_162 16200001 16300000
 AP008217_163 16300001 16400000
 AP008217_164 16400001 16500000
 AP008217_165 16500001 16600000
 AP008217_166 16600001 16700000
 AP008217_167 16700001 16800000
 AP008217_168 16800001 16900000
 AP008217_169 16900001 17000000
 AP008217_170 17000001 17100000
 AP008217_171 17100001 17200000
 AP008217_172 17200001 17300000
 AP008217_173 17300001 17400000
 AP008217_174 17400001 17500000
 AP008217_175 17500001 17600000
 AP008217_176 17600001 17700000
 AP008217_177 17700001 17800000
 AP008217_178 17800001 17900000
 AP008217_179 17900001 18000000
 AP008217_180 18000001 18100000
 AP008217_181 18100001 18200000
 AP008217_182 18200001 18300000
 AP008217_183 18300001 18400000
 AP008217_184 18400001 18500000
 AP008217_185 18500001 18600000
 AP008217_186 18600001 18700000
 AP008217_187 18700001 18800000
 AP008217_188 18800001 18900000
 AP008217_189 18900001 19000000
 AP008217_190 19000001 19100000
 AP008217_191 19100001 19200000
 AP008217_192 19200001 19300000
 AP008217_193 19300001 19400000
 AP008217_194 19400001 19500000
 AP008217_195 19500001 19600000
 AP008217_196 19600001 19700000

Alignment Scores:

Pred. No.: 1,73e-11 Length: 110000
 Score: 224.50 Matches: 51
 Percent Similarity: 65.7% Conservative: 16
 Best Local Similarity: 50.0% Mismatches: 24
 Query Match: 42.1% Indels: 11
 Gaps: 3

US-10-715-129-2 (1-106) x AP008217_078 (1-110000)

QY 3 ValG|A|ArgA|en|a|n|a|g|y|s|A|e|n|e|t|S|e|r|P|e|h|a|g|S|e|r|P|r|o|g|l|y|s|S|e|r|y|e|G|l|n| 22
 Db 58122 GTAGGTCTAGGCAGAGAGAGAAC---CTCTGATACCAGACCAACGGTTCAAGAGG 58066

QY 23 GluLeuArgArgSerphe-----SerAlaGlnLyArgMetMetIleProAlaAsn 39
 Db 58065 CATCTCAGCATGACATGACCAAAATGCTGCCGAAGAGAGACCCCGGACATGCAAT 58006
 QY 40 TyrPheSerLeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeu 59
 Db 58005 TACTTCAGTATCGAGCGCTTCCTCGTCTGCTCTTCTCCATGCATGTCATGCTCATCTT 57946
 QY 60 ProLeuValLeuProProLeuProProProPhenMetLeuLeuValProIleGly 79
 Db 57945 CCATTGTCCTTCCCTCCATGCTCCGCCCATGCTGCTGCTGCTGCAGCTTCG 57886
 QY 80 IleMetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsenAlaAsnThr 99
 Db 57885 CTCCTCATCTGCTGCTGCTTGTGCTGCTGCTCATGCCA-----ACG 57847
 QY 100 AspVal 101
 Db 57846 GATGTG 57841
 RESULT 13
 AC137924
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) chromosome 11 clone
 OSUNBA0095X08, complete sequence.
 AC137924
 VERSION
 AC137924.3 GI:46849661
 KEYWORDS
 HTG.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,C., Yuan,Q., Ouyang,S., Liu,J., Wang,A., Maiti,R., Jin,H.,
 Zhu,W., Hamilton,J., Jones,K., Tallon,L., Feldblyum,T., Teltrin,T.,
 Berra,J., Kim,M., Jin,S., Fadrosh,D., Vuong,H., Overton II,L.,
 Reardon,M., Weaver,B., Johri,S., Lewis,M., Utechtack,T., Van
 Aken,S., Wortman,J., Haas,B., Koo,H., Ziemann,V., Hsiao,J.,
 Iobst,S., de Vazelle,A., White,O., Salzberg,S. and Fraser,C.
 Unpublished
 2 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,R.
 Direct Submission
 Submitted (04-DEC-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 3 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,R.
 Direct Submission
 Submitted (10-APR-2003) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 4 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,R.
 Direct Submission
 Submitted (29-APR-2004) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 5 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,R.
 Direct Submission
 Submitted (19-APR-2005) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org
 6 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,R.
 Direct Submission
 Submitted (26-APR-2005) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org
 7 (bases 1 to 180999)
 REFERENCE
 AUTHORS
 Buell,R.
 Direct Submission
 Submitted (27-APR-2005) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org
 8 (bases 1 to 180999)
 REFERENCE


```

repeat_region /rpl_type=tandem
                /rpl_unit="cgs"
                /note="microsatellite"

Alignment Scores:
Pred. No.: 2.89e-11 Length: 180999
Score: 224.50 Matches: 51
Percent Similarity: 65.7% Conservative: 16
Best Local Similarity: 50.0% Mismatches: 24
Query Match: 42.1% Indels: 11
DB: 15 Gaps: 3

US-10-715-129-2 (1-106) x AC137924 (1-180999)

Qy 3 ValGlyArgAsnMetSerPheArgSerProGlyuSerLyGln 22
    |||||
Db 139826 GTTAGTCTTACGACAGAGAGAGAC--CTCTGTATACCCAGACCAAGGTTCAAGAGG 139882
    |||||

Qy 23 GluLeuArgArgSerPhe-----SerIaGlnArgMetMetIleProAlaAsn 39
    |||||
Db 139883 CATCTGAGCATGACATGACCAATGCTGCCAGAAAGAACCCGCCGACATCAAT 139942
    |||||

Qy 40 TyrPheSerLeuGlnSerLeuPheLeuValGlyLeuThraIaSerLeuLeu 59
    |||||
Db 139943 TACTTCAGTATCGAGCGGCTTCTGCTGCTGCTTCTTCCATCATGCTCATCTT 140002
    |||||

Qy 60 ProLeuValLeuProPoleuProProProPheMetLeuLeuValProIleGly 79
    |||||
Db 140003 CCATGAGTGTCTCCCGCATTTGCCCGCCGATGCTGCTGCTGCTGCTGCTGC 140062
    |||||

Qy 80 IleMetValLeuLeuValIleValIaPheMetProSerSerIaSerAlaAsnThr 99
    |||||
Db 140063 CTGCTCATCTCGCTGCTGCTGCTGCTGCTGCTTATGCA-----ACG 140101
    |||||

Qy 100 AspVal 101
    |||||
Db 140102 GATGTG 140107

RESULT 14
AK102641 1302 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J033100F02, full
DEFINITION insert sequence.
ACCESSION AK102641.1 GI:32987850
VERSION AK102641.1
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Bakayote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzeae; Oryza.

1 The Rice Full-length cDNA Consortium, National Institute of
  Agricultural Sciences Rice Full-length cDNA Project Team,
  Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
  Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
  Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
  Ohtsuki,T., Shishiki,T., Foundation of Advancement of International
  Science Genome Sequencing & Analysis Group; Ohtomo,Y., Murakami,K.,
  Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
  Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
  Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J.,
  Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Mura,J.,
  Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matubara,K., RIKEN:
  Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
  Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
  Kagawa,I., Kondo,S., Kono,H., Miyazaki,A., Otsu,N., Ota,Y.,
  Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
  Yoshino,M. and Hayashizaki,Y.
  Collection, mapping, and annotation of over 28,000 cDNA clones from
  japonica rice
  Science 301 (5631), 376-379 (2003)

```

```

PUBMED 12869764
2 (bases 1 to 1302)
REFERENCE Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
AUTHORS Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashidume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hirooka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Kono,H., Koude,M.,
Koya,S., Kurihara,C., Kurosaki,C., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matubara,K., Matsuyama,T., Mura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niihara,J., Nishii,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Otsu,N., Ota,Y., Ohtomo,Y., Ryu,R., Satoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) Shoichi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:kikuchi@nias.affrc.go.jp,
Tel:81-29-858-7007, Fax:81-29-858-7007)
COMMENT This clone is one of the 28k full-length cDNA clones from japonica
rice.
URL: http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ohtomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Mura,J.,
Mizuno,K., Narikawa,R., Niihara,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashidume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hirooka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Kono,H., Koude,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Ohtomo,Y.,
Oka,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.
FEATURES
source Location/Qualifiers
1. 1302
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033100F02"

ORIGIN
Alignment Scores:
Pred. No.: 6.78e-12 Length: 1302
Score: 209.00 Matches: 40
Percent Similarity: 90.0% Conservative: 14
Best Local Similarity: 66.7% Mismatches: 6
Query Match: 39.2% Indels: 0
DB: 15 Gaps: 0

```

[illegible]

FEATURES

source

location/Qualifiers

1. .46422

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="japonica"

/cultivar="Nipponbare"

/sub_species="japonica"

/db_xref="taxon:39947"

/chromosome="4"

/clone="OSUNBa0058603"

/clone_1id="CUGI-OSUNBa"

/complement (1144. .2100)

/gene="OSUNBa0058603.1"

/complement (1144. .2100)

/gene="OSUNBa0058603.1"

/codon_start=1

/protein_id="CAE02427.2"

/db_xref="GI:38344583"

/db_xref="UniProt/TREMBL:O7P909"

/translation="MHPTPTNTIIGVTYVLISIVSLGICLCHSLNFPULLKRR
RNYQVANDQELSYNGPMLTRITLILVALMWGEVRLITVNEBGRISQTVQAN
KCTPVNLGFBPGLFLFLAFLISALQNDVBLKRMQRICAVFMICPSLIW
EACVPIGPIIASNDQTSKVAKYWASVVDADVCTYELISIFGTVTYLY
VIFPGQILSIVINKGRRIIRYVYLIFANGILLPRAMIFGVLPMWPGIVHBSLVFS
PLVIMAMIGVILVYFPVAFETEVNQHIEIQTSHSL"

/complement (6127. .6681)

/gene="OSUNBa0058603.2"

/complement (6127. .6681)

/gene="OSUNBa0058603.2"

/codon_start=1

/protein_id="CADA1044.1"

/db_xref="GI:21742878"

/db_xref="GOA:Q7XUS1"

/db_xref="InterPro:IPR001471"

/db_xref="UniProt/TREMBL:Q7XUS1"

/translation="MASAAAAAGCGCYRGGRRRKGKWSRIROGPKTIRIWSFE
SABMAVAHYVAALRIKGRDQQLPSSVDRLPRPASRPDIDIAAAEAADRVREP
ALVAVDSEMGEGFLAVRNAGLEVBVEQQLGSSDEFEVDSPLIMEMALMLDPP
VMVAVDSEMGEGFLAVRNAGLEVBVEQQLGSSDEFEVDSPLIMEMALMLDPP

/complement (17958. .18485)

/gene="OSUNBa0058603.3"

/complement (17958. .18485)

/gene="OSUNBa0058603.3"

/codon_start=1

/protein_id="CAE02428.1"

/db_xref="GI:32482926"

/db_xref="GOA:Q7XRM1"

/db_xref="InterPro:IPR007087"

/db_xref="UniProt/TREMBL:Q7XRM1"

/translation="MEBGMVEDAAAAGEBELLTSLSPSPRRFOALSCCYCPRPK
RSQALGQNAHKLQRIALRAGREAAASLAAAAAASSGDQCCQRTTAAA
AYLGGESAPPAARAADLDGAGVWGAGKRGKRAHHRMLMGQSSGSSAAGRG
NGELADEMIDSLTL"

24188. .29072

/gene="OSUNBa0058603.4"

/complement (24188. .24309,25603. .25678,25778. .25848,25983. .26476
26560. .26720,26810. .27013,27773. .28007,28477. .29072)

/gene="OSUNBa0058603.4"

/codon_start=1

/protein_id="CAE02429.1"

/db_xref="GI:32482927"

/db_xref="UniProt/TREMBL:Q7XRM0"

/translation="MORGRILERSGRALADAGGGGDDDDRAKPRPRVPLASVIV
EALVDSLOKLCSSLEPIRLRVSEBVRALAKIPATPRIGRSSPKKIDVYLG
NLOLQPSRLSLPTFGKVGEGGALIHVLLDANGRVVTVGSPSKADIVLYEG
DNKEQDBDTBEPEPHIYKEREGKRPILGDIQVTLKEVGTIGELIFDNSMIR
SRKRLGLRVSSGFCBGRVVEATBAFTVVDHGEIYKQYIPALKODVRLKLGK
DGAKRLKLSNGIYVHFLOLIVRDOOKRTIIGSNNSNMWESLVEHATCVLSGK
HIYSSDARSVGAIPANNIYEFGLIADDOYISAEINLSENGRLPADTLVQAYDWM
VYEDGKELRLFKQKKSVTTRSDTAKASTYPSYSGTSHKQLTGSPVNIQBSMS
SMASDGRMNSNGSOARVYANPQDIOSTIMPYDMSLRPEQFYSSTIOQASRS
NMLAGPTQOONFBSALGSGMQSPPLNPPDMSRLQENRGQDYLMBZIRVASHZL
LENEDEDQOQMRILSMGSSANMNGDSEFPMSPAAFVYEDDRAPSGAIVVGL

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 30, 2006, 06:39:18 ; Search time 493 Seconds
(without alignments)
1432.977 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533
Sequence: 1 MDVGRNRRKMKSRSPRSPEKS.....LAFWPSHSNANTDYTCNFM 106

Scoring table:

| | |
|---------------------------|--|
| BLOSUM62 | |
| Xgapop 10.0 , Xgapext 0.5 | |
| Ygapop 10.0 , Ygapext 0.5 | |
| Fgapop 6.0 , Fgapext 7.0 | |
| Delop 6.0 , Delext 7.0 | |

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10715129/rnat.27012006.154744.14557/app.query.fasta_1.263
-DB=N Geneseg -OFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdd -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USERS=US10715129 @CGN 1.1 1096 @rnat.27012006.154744.14557 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21:*

- 1: genesegq1980s:*
- 2: genesegq1990s:*
- 3: genesegq2000s:*
- 4: genesegq2001as:*
- 5: genesegq2001bs:*
- 6: genesegq2002as:*
- 7: genesegq2002bs:*
- 8: genesegq2003as:*
- 9: genesegq2003bs:*
- 10: genesegq2003cs:*
- 11: genesegq2003ds:*
- 12: genesegq2004as:*
- 13: genesegq2004bs:*
- 14: genesegq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|--------|--------------------|
| 1 | 533 | 100.0 | 732 14 | ADZ67285 |
| 2 | 327 | 61.4 | 336 6 | ABZ13462 Arabidops |
| 3 | 327 | 61.4 | 336 8 | ADA67786 Arabidops |
| 4 | 327 | 61.4 | 1005 3 | AAC53386 Arabidops |

| | | | | |
|----|-------|------|-----------|------------|
| 5 | 235 | 44.1 | 650 3 | AAC51918 |
| 6 | 224.5 | 42.1 | 1137 14 | ABE67601 |
| 7 | 215 | 40.3 | 298 11 | ACL31088 |
| 8 | 197.5 | 37.1 | 978 11 | ACL29210 |
| 9 | 195 | 36.6 | 665 13 | ADK49328 |
| 10 | 193 | 36.2 | 990 13 | ADK49328 |
| 11 | 193 | 36.2 | 1157 13 | ADK49328 |
| 12 | 181 | 34.0 | 1055 13 | ADT18995 |
| 13 | 165.5 | 31.1 | 939 13 | ADT17295 |
| 14 | 154.5 | 29.0 | 936 13 | ADK37027 |
| 15 | 149.5 | 28.0 | 366 13 | ADK49328 |
| 16 | 144 | 27.0 | 577 13 | ADK49328 |
| 17 | 144 | 27.0 | 2000 6 | ABZ16334 |
| 18 | 144 | 27.0 | 2000 8 | ADA68185 |
| 19 | 140 | 26.3 | 473 8 | ACC60011 |
| 20 | 132 | 24.8 | 647 13 | ADT17651 |
| 21 | 121.5 | 22.8 | 501 3 | AAC54264 |
| 22 | 115 | 21.6 | 529 3 | AAC53305 |
| 23 | 106 | 19.9 | 1863 5 | AAS80665 |
| 24 | 105 | 19.7 | 2010 13 | ADU01858 |
| 25 | 104 | 19.5 | 202251 11 | ACN44504 |
| 26 | 99.5 | 18.7 | 110000 8 | ABX16390_1 |
| 27 | 99 | 18.6 | 43229 11 | ACN44192 |
| 28 | 99 | 18.6 | 235962 14 | ADZ12926 |
| 29 | 98.5 | 18.5 | 1197 5 | AAS89331 |
| 30 | 98.5 | 18.5 | 4128 5 | AAS88553 |
| 31 | 97 | 18.2 | 110000 8 | ABX16390_3 |
| 32 | 96.5 | 18.1 | 575 4 | AA117342 |
| 33 | 96.5 | 18.1 | 575 4 | ABA62258 |
| 34 | 96.5 | 18.1 | 575 4 | AA142236 |
| 35 | 96.5 | 18.1 | 575 4 | ABA29604 |
| 36 | 96.5 | 18.1 | 575 4 | AAK36486 |
| 37 | 96.5 | 18.1 | 575 4 | AAK10591 |
| 38 | 96.5 | 18.1 | 575 4 | ABS36136 |
| 39 | 96.5 | 18.1 | 575 6 | ABS10483 |
| 40 | 96.5 | 18.1 | 30306 11 | ACN44196 |
| 41 | 96 | 18.0 | 287 4 | AA126561 |
| 42 | 96 | 18.0 | 287 4 | ABA74794 |
| 43 | 96 | 18.0 | 287 4 | AA155319 |
| 44 | 96 | 18.0 | 287 4 | ABA39495 |
| 45 | 96 | 18.0 | 287 4 | AAK49450 |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| ADZ67285 | |
| ID | ADZ67285 standard; cDNA; 732 BP. |
| AC | ADZ67285; |
| DT | 28-JUL-2005 (first entry) |
| DE | Arabidopsis thaliana auxin-inducible gene ARGOs. |
| OS | Arabidopsis thaliana. |
| PH | Key |
| FT | CDS |
| FT | Location/Qualifiers |
| FT | 388..708 |
| FT | /*tag= a |
| FT | /product= "ARGOS protein" |
| PN | US2005106793-A1. |
| XX | 19-MAY-2005. |
| PD | 18-NOV-2003; 2003US-00715129. |
| XX | 18-NOV-2003; 2003US-00715129. |
| PR | 18-NOV-2003; 2003US-00715129. |
| XX | (TEMA-) TEMASEK LIFE SCI LAB. |
| PA | |

XX Hu Y, Xie Q, Chua N;
 XX WPI; 2005-354827/36.
 DR P-PSDB; AD267286.
 DR GENBANK; AY305869.
 XX
 PT New auxin-inducible gene ARGOS from Arabidopsis, useful for regulating
 PT organ development including size control in transformed plants and plant
 PT cells.
 PS Claim 1; SEQ ID NO 1; 22pp; English.
 XX
 CC The invention relates to an isolated nucleic acid, i.e. auxin-inducible
 CC gene, ARGOS, from Arabidopsis, comprising fully defined 732 base pair
 CC sequence given in the specification, that encodes the amino acid sequence
 CC comprising fully defined 106 amino acids also given in the specification,
 CC or comprising their antisense nucleotide sequences. The nucleic acid is
 CC useful for regulating organ development including size control in
 CC transformed plants and plant cells. This sequence corresponds to the cDNA
 CC for the novel auxin-inducible gene ARGOS.
 XX
 SQ Sequence 732 BP; 199 A; 170 C; 125 G; 238 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,1e-52 Length: 732
 Score: 533.00 Matches: 106
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0
 US-10-715-129-2 (1-106) x AD267285 (1-732)

QY 1 MetAspValGlyArgAsnAsnArgLysAsnMetSerPheArgSerProGlnLysSer 20
 DB 388 ATGGACCTCGAAGACAAACCGGAAAAACATGAGCTTCCAGAGTCCCGGAGAAAAAGC 447
 QY 21 LysGlnGlnLeuLeuArgSerPheSerAlaGlnLysArgMetMetIleProAlaAsnTyr 40
 DB 448 AAGCAAGAGTTACCGCGAGTTTCGCGCGAGAAAGAGAGATGCCGCGCAATTAT 507
 QY 41 PheSerLeuGlnSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuLeuPro 60
 DB 508 TTGAGTTTAAAGTCTGCTGCTTCTATGCTGTTAAGCATCTCTGTTAATCTCCG 567
 QY 61 LeuValLeuProProLeuProProProPheMetLeuLeuValProIleGlyIle 80
 DB 568 TTAGTTTTCGCGCGGATACCTCCGCTCCGTTATGCTGCTATGTTCCCATGGGATT 627
 QY 81 MetValLeuLeuValValLeuAlaPheMetProSerSerHisSerAsnAlaAsnThrAsp 100
 DB 628 ATGGTTTACTCGTCTGCTTCTGCTTCAATGCTTCTCTCAATCAATACAGAT 687
 QY 101 ValThrCysAsnPheMet 106
 DB 688 GTAACTTGCAATTTCATG 705
 RESULT 2
 AB213462 standard; DNA; 336 BP.
 XX
 XX AB213462;
 XX
 DT 21-JAN-2003 (first entry)
 XX
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1267.
 XX
 KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200216655-A2.

XX
 PD 28-FEB-2002.
 XX
 XX 24-AUG-2001; 2001MO-US026685.
 PF
 XX 24-AUG-2000; 2000US-0227866P.
 PR 26-JAN-2001; 2001US-0264647P.
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Harper UF, Kreps J, Wang X, Zhu T;
 XX
 DR WPI; 2002-304127/34.
 XX
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 14; SEQ ID NO 1267; 577pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 SQ Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,26e-28 Length: 336
 Score: 327.00 Matches: 72
 Percent Similarity: 82.2% Conservative: 11
 Best Local Similarity: 71.3% Mismatches: 12
 Query Match: 61.4% Indels: 6
 DB: 6 Gaps: 3
 US-10-715-129-2 (1-106) x AB213462 (1-336)

QY 1 MetAspVal--GlyArgAsnAsnArgLysAsnMetSerPheArgSerPhe----- 16
 DB 1 ATGGACGTGAGAGAGATCATTAACCGGAAAAACAGAGTTTGTGCTGATCCAGCT 60
 QY 17 ProGlnLysSerLysGlnGlnLysArgSerPheSerAlaGlnLys-----ArgMet 34
 DB 61 CCGATTATGGGGAAGCAAGATGTTTTCGAGCATTTGCTCGCGAAGCAAGTCCAGGAGG 120
 QY 35 MetIleProAlaAsnTyrPheSerLeuGlnSerLeuPheLeuValGlyLeuThrAla 54
 DB 121 CTAAATATCAGCAGGATTAATCTCAAGTTTGAATATCAATGTTGCTGTTGATCTACAGCA 180
 QY 55 SerLeuLeuLeuLeuProLeuValLeuProProLeuProProPheMetLeuLeu 74
 DB 181 TCTCTCTTGAATCTTACCGTTGATTTCCACCATTCGCTCTCTCTTATATGCTGCTT 240
 QY 75 LeuValProIleGlyIleMetValLeuValValLeuAlaPheMetProSerSerHis 94
 DB 241 TTGATTCCTATGAGGATTAATGTTTGTCTTATGCTTCTTCAATCCCTTCTTCAAT 300
 QY 95 Ser 95
 DB 301 TCC 303
 RESULT 3
 ADA67786 standard; DNA; 336 BP.
 ID ADA67786

XX ADA67786;
AC
XX 20-NOV-2003 (first entry)
DT
XX Arabidopsis thaliana gene, SEQ ID 23.
DE
XX Plant; bacterial infection; fungal infection; viral infection; ds.
KM
XX Arabidopsis thaliana.
OS
XX WO2003000898-A1.
PN
XX 03-JAN-2003.
PD
XX 22-JUN-2001; 2001WO-IB001105.
PF
XX 22-JUN-2001; 2001WO-IB001105.
PR
XX 22-JUN-2001; 2001WO-IB001105.
PS (SYGN) SYNGENTA PARTICIPATIONS AG.
PA
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
DR
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 23; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 336 BP; 72 A; 74 C; 68 G; 122 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,266-28 Length: 336
Score: 327.00 Matches: 72
Percent Similarity: 82.2% Conservative: 11
Best Local Similarity: 71.3% Mismatches: 12
Query Match: 61.4% Indels: 6
DB: Gaps: 3

US-10-715-129-2 (1-106) x ADA67786 (1-336)
QY 1 MetAapVal---GIATgsmAaAnAgiysAaMteSerPheArGserSer----- 16
Db 1 ATGAGCGTGAAGAGATCATTAACCGGAAAAACAGATTTTCGTGCTTCAGCTCAGCT 60
QY 17 ProGluIuyseryGInGInIuLeuArGserPheSerIaGInIys-----ArgMet 34
Db 61 CCGATTATGGGGAAGCAAGAAATTGTTTCGACATTGCTGTCGAGAACAGTCCAAAGAGG 120
QY 35 MetIleProIaAaNYrPheSerIeugIuSerIeupheLeuLeuValGlyLeuThrIa 54
Db 121 CTATATATACGAGATTAATCTTCAGTTAGATCATGCTGTGCTTGTCTCAGCA 180
QY 55 SerLeuLeuIleuProLeuValIuProLeuProLeuProProProPheMetIleu 74
Db 181 TCTCTCTGATCTTAACCGTGAATCTTCCACCATTCCTCTCTCTTATATGCTGCTT 240
QY 75 LeuValProIleGlyIleMetValIuLeuValIalleuAlaPheMetProSerSerHis 94

Db 241 TTGATTCCTATTCGATTAAGCATTTTCCTATATGCTTTCATGCTTCTTAT 300
QY 95 Ser 95
Db 301 TCC 303

RESULT 4
AAC53386
ID AAC53386 standard; DNA, 1005 BP.
AC
XX AAC53386;
AC
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 74346.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KM promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145145P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-715-129-2 (1-106) x AAC53386 (1-1005)

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

1005

72

11

12

6

3

Qy 1 MetAspVal---(G)ArgAsnAsnArgIysAsnMetSerPheArgSerSer----- 16
Db 617 ATGGACGGAGACGAGATCATACCGGAAACAGACTTTTCGTGGTTCACCTCAGCT 676
Qy 17 ProGluIysSerIysGInGluLeuArgSerPheSerAlaGlnIys-----ArgMet 34
Db 677 CCGATTATGGGAGGACCAAGATTGTTTCGACATGTGTCGCAAGACAGTCCAAGAGG 736
Qy 35 MetIleProAlaAsnIYrPheSerIeuGInSerIeuPheIleuValGlyLeuThrAla 54
Db 737 CTATATCAGCGAGACTTACTTCAGTTAGAAATCAATGCTTGCTTGCTGCTCAGCA 796
Qy 55 SerIeuIleuIleuProIeuValIeuProIleuProIleuProIleuMetIleu 74
Db 797 TCTCTCTGATCTTACCGTGAATCTTCACCATTCCTCCTCTTATATGCTGCTT 856
Qy 75 LeuValProIleGlyIleMetValIeuIeuValIleuAlaPheMetProSerSerHis 94
Db 857 TTGATTCCTATGGGATTAATGCTTTTCTTATGCTTCTGCTTCTCATGCTTCTTAT 916
Qy 95 Ser 95
Db 917 TCC 919
RESULT 5
AAC51918
ID AAC51918 standard; DNA; 650 BP.
XX AAC51918;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 70014.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131443P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134916P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.
PR 18-JUN-1999; 99US-0139483P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140658P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145090P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.

| | | |
|---|---|------------------|
| PR | 25-OCT-1999; | 99US-0161406P. |
| PR | 26-OCT-1999; | 99US-0161359P. |
| PR | 26-OCT-1999; | 99US-0161360P. |
| PR | 26-OCT-1999; | 99US-0161361P. |
| PR | 28-OCT-1999; | 99US-0161920P. |
| PR | 28-OCT-1999; | 99US-0161922P. |
| PR | 28-OCT-1999; | 99US-0161932P. |
| PR | 29-OCT-1999; | 99US-0162142P. |
| Alignment Scores: | | |
| Pred. No.: | 1,35e-17 | Length: 650 |
| Score: | 235.00 | Matches: 53 |
| Percent Similarity: | 71.7% | Conservative: 13 |
| Best Local Similarity: | 57.6% | Mismatches: 14 |
| Query Match: | 44.1% | Indels: 12 |
| DB: | 3 | Gaps: 3 |
| US-10-715-129--2 (1-106) x AAC51918 (1-650) | | |
| Oy | 13 PheArgSerSerProGlnYSerIysGlnJleuAry-----ArgSerPheSerAla | 30 |
| Dd | ::::TATCTATCGAGTCCTGTAATAACAAGAATCAACATTTCAGACTTGAGGGAATGGTTGCCA | 37 |
| Oy | 31 GlnYArgMet-----MetLleProAlaSerTyPheSerLeu | 43 |
| Dd | 378 GATCCAACTGTAGAAAACAATCCTATTATACAGGTGATTCCTGT-----TTTTTA | 428 |
| Oy | 44 GluSerIserPheLeuLeuValAlGlyLeuThrAlaSerLeuLeuIleLeuProLeuValLeu | 63 |
| Dd | 429 GACTCAATGGTTGTGCTTTGGTCTCACAGCATCTCTTGATCTTACCCTTAATCTT | 488 |
| Oy | 64 ProProLeuProBofProBofProPhemeIleuLeuValProIleGylIleMetValLeu | 83 |
| Dd | 489 CCACCAATGGCTCTCTCTCTTTTATAGCTGCTTTTGATTCATGGATTGGATTATGGTTTG | 548 |
| Oy | 84 LeuValValLeuAlaPheMetProSerSerHisSer | 95 |
| Dd | 549 CTTATGGTCTTGCTTTTCATAGCCTTCTCTTAATTC | 584 |
| RESULT 6 | | |
| AEB67601 | AEB67601 standard; DNA; 1137 BP. | |
| XX | AEB67601; | |
| AC | 22-SEP-2005 (first entry) | |
| DT | | |
| XX | Rice genome derived DNA sequence, SEQ ID 2746. | |
| DE | | |
| XX | transcription; gene regulation; transgenic plant; RNA interference; | |
| KW | transformation; antibody; ds. | |
| OS | Oryza sp. | |
| PN | JP2005185101-A. | |
| PD | 14-JUL-2005. | |
| PF | 11-DEC-2002; 2002JP-00383870. | |
| PR | 30-MAY-2002; 2002JP-00203269. | |
| PA | (DOKU-) DOKURITSU GYOSEI HOJIN NOGO SEIBUTSU SH. | |
| PA | (SEIB-) SEIBUTSUKAI TOKUTEI SANGYO GIJITSU. | |
| PA | (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKYUSH. | |
| PA | (KOKU-) ZH KOKUSAI KAGAKU SHINKO ZAIDAN. | |
| PI | Kikuchi H, Hayashizaki Y, Ootomo Y, Matsubara K, Murakami K; | |
| PI | Kishimoto N, Sato K, Negata T, Kawakami N, Yarakai J, Ishikawa M; | |
| PI | Doi K, Kawai J; | |
| DR | WPI; 2005-566181/58. | |

XX DE Rice abiotic stress responsive polynucleotide SEQ ID NO:3166.
 XX KM 88; abiotic stress tolerance; transgenic plant; plant; cereal;
 XX KM agriculture.
 XX OS Oryza sativa.
 XX PN MO2003008540-A2.
 XX PD 30-JAN-2003.
 XX PF 21-JUN-2002; 2002MO-US019668.
 XX PR 22-JUN-2001; 2001US-0300112P.
 XX PR 24-AUG-2001; 2001US-0314662P.
 XX PR 26-SEP-2001; 2001US-0325277P.
 XX PR 21-NOV-2001; 2001US-0332132P.
 XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PI Kieps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
 XX PI Moughamer T, Provart N, Riecke D, Zhu T;
 XX DR WPL; 2003-248011/24.
 XX PT New stress-responsive nucleic acid, useful for altering the
 XX PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 XX PT stress, salt stress or osmotic stress.
 XX PS Claim 1; SEQ ID NO 3166; 89pp; English.
 XX CC The invention relates to novel abiotic stress responsive polynucleotides
 XX CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 XX CC cells, and plants containing such polynucleotides. Also disclosed are
 XX CC methods for using the polynucleotides and polypeptides to alter the
 XX CC responsiveness of a plant to abiotic stress. The invention is useful in
 XX CC agriculture. The nucleic acid is useful for determining whether a test
 XX CC plant has been exposed to an abiotic stress condition. It is also useful
 XX CC for selecting an agent that alters abiotic stress regulated
 XX CC polynucleotide expression in a plant cell, and to identify a homolog or
 XX CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 XX CC molecule and the polypeptide encoded by it are useful in altering the
 XX CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 XX CC stress, osmotic stress or any of their combinations. The present sequence
 XX CC is used in the exemplification of the invention.
 XX SQ Sequence 978 BP; 253 A; 269 C; 210 G; 246 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,08e-13 Length: 978
 Score: 197.50 Matches: 43
 Percent Similarity: 61.2% Conservaive: 20
 Best Local Similarity: 41.7% Mismatches: 23
 Query Match: 37.1% Indels: 17
 DB: 11 Gaps: 2
 US-10-715-129-2 (1-106) x ACL29210 (1-978)
 QY 1 MetApValGtAgaAmAmArUlyAaMserPheArGserSerProGluLysSer 20
 Db 655 ATGACAGAGAGGTGAGACATCATCATCATCACACCAAGGAACAAACACACAGAG 714
 QY 21 LygdlInuLeuAaGArGserPheSerAlaInLyArGMeTtleProAlaAn--- 39
 Db 715 AAGCAGACAGACGCGCGG-----CGGCTGATGAACAATGCACCAAC 756
 QY 40 -----TyrPheSerLeuGluSerLeuPheLeuLeu 49
 Db 757 GCGTGTCTTCTTACTGCGGACAGAGTGTCTTCTTACACAGGAGGCATCTGTCTG 816
 QY 50 ValGlyLeuThrAlaSerLeuLeuIleuProLeuValLeuProProLeuProPro 69
 ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 817 GCATGCGTACCGGTGTGCTGTGTGTGTCGCCCTCATCTGCGCGCGCGCGCG 876
 QY 70 ProPheMetLeuLeuLeuValProIleGlyIleMetValLeuValIleAlaPhe 89
 Db 877 CCAGCGTGTGTGCTGTGTGTCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 936
 QY 90 MetProSer 92
 Db 937 ATGCCCACT 945
 RESULT 9
 ADX49328
 ID ADX49328 standard; cDNA; 665 BP.
 XX AC ADX49328;
 XX XX
 XX DT 21-APR-2005 (first entry)
 XX XX
 XX DE Plant full length insert polynucleotide seqid 24068.
 XX XX
 KM plant protectant; plant growth regulator; gene therapy; plant;
 KM recombinant DNA construct; physical array; plant breeding marker;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KM extreme osmotic condition; pathogen tolerance; pest tolerance;
 KM growth rate; cell cycle pathway; disease resistance;
 KM galactomanan production; lignin production; plant growth regulator;
 KM yield; plant growth; plant development; seed oil; protein yield;
 KM protein content; gene; ss.
 XX OS Unidentified.
 XX OS
 XX PN US2004034888-A1.
 XX PD 19-FEB-2004.
 XX PF 28-APR-2003; 2003US-00425114.
 XX PR 06-MAY-1999; 99US-00304517.
 XX PR 05-NOV-2001; 2001US-00985678.
 XX PA (LIU/) LIU J.
 XX PA (ZHOU/) ZHOU Y.
 XX PA (KOVA/) KOVALIC D K.
 XX PA (SCORE/) SCREEN S E.
 XX PA (TABAS/) TABASKA J B.
 XX PA (CAO/) CAO Y.
 XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;
 XX PI WPL; 2004-180133/17.
 XX DR
 XX PT New recombinant DNA construct, useful for improving plant tolerance to
 XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 XX PT pests, for conferring increased resistance to plant disease, or for
 XX PT improving yield.
 XX PS Claim 1; SEQ ID NO 24068; 15pp; English.
 XX XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomanan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or

CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biochemistry and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC Seqdata.uspto.gov/sequence.html?docID=20040216190.

XX SQ Sequence 1055 BP; 218 A; 331 C; 268 G; 238 T; 0 U; 0 Other;

Alignment Scores:

Score: 4.57e-11 Length: 1055
 181.00 Matches: 45
 Percent Similarity: 63.1% Conservative: 20
 Best Local Similarity: 43.7% Mismatches: 33
 Query Match: 34.0% Indels: 5
 DB: 13 Gaps: 2

US-10-715-129-2 (1-106) x ADT18995 (1-1055)

QY 5 ArgAhnAnArgLYsAsmMet-SerPhaArgSerProGluYsSerLYsGInGluLe 24
 Db 526 AGGAGCGCGCGGAGGACGTCGTGCGCATCGCCGCGTGATCCGCGCGGAGCA 585
 QY 24 uArgARSerPheserAlaGlnLYs-----ArgMetMetIleProAlaAsnTYrph 41
 Db 586 GAAAGCAGCGTACGAGGAGGAGGACCCGCGTGTGACAGCTGACGCGATGCTT 645
 QY 41 e---SerLeuGluSerLeuPhelauValGlyLeuThrAlaSerLeuLeuIleuPr 60
 Db 646 CGGCAAGTACTTCTGCTCTCTGCTCATATGTCACCCGCGCGTGTGATTTCTCC 705
 QY 60 OleuValleuProProleuProProProPhemeuLeuValProIleGlyI 80
 Db 706 GCTGTCTCTGACCGCGTCTGCTCCGCGCGCTGATGCTGCTGCGGCGGCAAT 765
 QY 80 emetValleuValleuValleuAlaPhemeProSerSerHisSerAlaAlaAsnThra 100
 Db 766 GCTGTGATGCTCTGCTGCTGCGTTCATGCGGAGAGTGGCGGCGGAGCGG 825
 QY 100 pValThr 102
 Db 826 CCCGACT 832

RESULT 13
 ADT17295/c
 ID ADT17295 standard; cDNA; 939 BP.

XX AC ADT17295;
 XX DT 13-JAN-2005 (first entry)
 XX DB Plant cDNA, Seq ID 2621.

XX KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
 KW drought tolerance; disease resistance; galactomannan production;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW lignin production; extreme osmotic condition tolerance;
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KW seed protein yield.
 OS Vitidiplantae.
 PN US2004216190-A1.
 XX 28-OCT-2004.
 XX 18-DEC-2003; 2003US-00739930.
 XX 28-APR-2003; 2003US-00424599.
 XX 28-APR-2003; 2003US-00425115.
 XX (KOVA/) KOVALIC D K.
 XX KOVALIC DK;
 XX WPI; 2004-757369/74.
 DR New recombinant DNA constructs useful in the field of biochemistry and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX Claim 1; SEQ ID NO 2621; 14pp; English.

PS The invention relates a recombinant DNA construct comprising a
 XX polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed
 CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biochemistry and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen
 CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant cDNA
 CC sequences of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC Seqdata.uspto.gov/sequence.html?docID=20040216190.

XX SQ Sequence 939 BP; 215 A; 263 C; 287 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Score: 2.43e-09 Length: 939

Score: 165.50 Matches: 44
 Percent Similarity: 60.6% Conservative: 19
 Best Local Similarity: 42.3% Mismatches: 38
 Query Match: 31.1% Indels: 4
 DB: 13 Gaps: 2

US-10-715-129-2 (1-106) x ADT17295 (1-939)

Qy 5 ArgAsnAsnArgLyAsnMetSerPhe-----ArgSerSerProGlyLysSerLyGln 22
 Db 433 CAGAAATGAGAGCGACCATGCTGCTGCAAGAAAGAGCCCGACGGCTGGCCAAAGGG 374
 Qy 23 GluLeuValArgSerPheSerAlaGlnLysArgMetMetIleProAlaAsnTyrPheSer 42
 Db 373 CACGACGCTGAGAAAGCTTCGCCCGCTGCCCAAGCTCCGCCCGGGGGGCTCAGC 314
 Qy 43 LeuGlnSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuIleLeuProLeuVal 62
 Db 313 GCGAGAGCGTTCCTGCTTCTGCGTCCGCTGCTGCTTATCGTCTGCCGCTGCTC 254
 Qy 63 LeuProProLeuProProProPheMetLeuLeuValProIleGlyIleMetVal 82
 Db 253 CTGCGCGCGCTGCCGCCCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
 Qy 83 LeuLeuValValLeuAla---PheMetProSerSerHisSerAlaAlaAsnThrAspVal 101
 Db 193 CTCCTCCCGCGCTGCCACCTTCATGCGCTGGATG--TCAGAGCATGCACTCTCCA 136
 Qy 102 ThrCysAsnPhe 105
 Db 135 ACTGTAACTAC 124

RESULT 14
 ADX37027/c
 ID ADX37027 standard; cDNA; 936 BP.

AC ADX37027;

DT 21-APR-2005 (first entry)

DE plant full length insert polynucleotide seqid 19847.

XX plant protectant; plant growth regulator; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/) LIU J.

XX (ZHOU/) ZHOU Y.
 XX (KOVA/) KOVALIC D K.
 XX (SCRE/) SCREEN S E.
 XX (TABA/) TABASKA J E.
 XX (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,
 XX WPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

XX Claim 1; SEQ ID NO 19847; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 936 BP; 247 A; 221 C; 234 G; 234 T; 0 U; 0 Other;

XX Alignment Scores:

| Pred. No.: | Length: | Score: | Matches: |
|------------------------|---------|--------|-----------------|
| 4.55e-08 | 936 | 154.50 | 33 |
| Percent Similarity: | | 60.0% | Conservative: 9 |
| Best Local Similarity: | | 47.1% | Mismatches: 17 |
| Query Match: | | 29.0% | Indels: 11 |
| DB: | | 13 | Gaps: 1 |

US-10-715-129-2 (1-106) x ADX37027 (1-936)

Qy 25 ArgArgSerPheSerHisGlnLysArgMetMetIleProAlaAsnTyrPheSerLeuGlu 44

Db 746 GCGAGGCGCGCGCTGATGACCAAGAGGTA----- 717

Qy 45 SerLeuPheLeuLeuValGlyLeuThrAlaSerLeuLeuIleLeuProLeuValLeuPro 64

Db 716 ---CTGATGATGTTGCTGGCGGAGACATGATCTCTGCTGCTGCTGCTGCTGCTGCTG 660

Qy 65 ProLeuProProProPheMetLeuLeuValProIleGlyIleMetValLeuLeu 84

Db 659 CCCCTCCCGCACACCGCTGTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

Qy 85 ValValLeuAlaPheMetProSerSerHis 94

Db 599 TTCTCCCTGGTTTCTTCCGCTCAACAC 570

RESULT 15
 ADR64525
 ID ADR64525 standard; cDNA; 366 BP.

XX ADR64525;

XX 02-DEC-2004 (first entry)

XX Cotton cDNA sequence, SEQ ID 5306.

XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
 KW drought tolerance; plant disease resistance; galactomannan; lignin;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW homologous recombination; extreme osmotic condition tolerance;
 KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
 KW stress resistance.

XX Gossypium hirsutum.

PN US2004181830-A1.
XX 16-SEP-2004.
XX 29-JAN-2004; 2004US-00767795.
XX 07-MAY-2001; 2001US-00849529.
PR 12-DEC-2001; 2001US-00021323.
XX (KOVA/) KOVALIC D K.
PA (ZHOU/) ZHOU Y.
PA (CAOY/) CAO Y.
XX
PI Kovalic DK, Zhou Y, Cao Y;
XX WPI; 2004-667718/65.
XX
XX New recombinant nucleic acid molecules and polypeptides from *Gossypium*
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX
PS Claim 1; SEQ ID NO 5306; 14pp; English.
XX
XX The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomanan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.
XX
SQ Sequence 366 BP; 65 A; 112 C; 71 G; 118 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4,94e-08 Length: 366
Score: 149.50 Matches: 33
Percent Similarity: 70.0% Conservative: 9
Best Local Similarity: 55.0% Mismatches: 14
Query Match: 28.0% Indels: 4
DB: 13 Gaps: 1
US-10-715-129-2 (1-106) x ADR64525 (1-366)
QY 37 Prok|a|e|n|t|y|r|p|e|-----S|e|r|i|e|n|g|u|-----S|e|r|i|e|n|p|h|e|u|e|u|a|l|g|l|y|e|u|t|h| 53
Db 97 C|C|A|T|G|C|A|C|A|T|T|T|C|A|G|A|G|T|G|A|G|C|T|T|A|A|G|C|C|G|A|T|C|A|T|G|C|T|T|G|C|C|G|T|T|C|G|C| 156
QY 53 r|A|S|e|r|i|e|u|e|u|l|e|u|e|u|P|r|o|l|e|g|l|y|l|l|e|u|e|u|a|l|e|u|a|l|a|P|h|e|u|e|t|P|r|o|s|e|r| 73
Db 157 T|G|T|C|T|T|C|T|T|T|G|T|G|G|C|T|C|T|G|T|T|T|C|C|C|G|C|A|T|T|G|C|G|C|C|T|C|G|T|C|T| 216

QY 73 u|e|u|e|u|a|l|P|r|o|l|e|g|l|y|l|l|e|u|e|u|a|l|e|u|a|l|a|P|h|e|u|e|t|P|r|o|s|e|r| 92
Db 217 T|C|T|T|A|T|C|C|A|G|T|T|G|A|T|G|A|T|G|C|T|G|C|T|C|T|T|T|C|C|G|C|A|T|T|T|C|A|C|C|G|C|T| 274

Search completed: January 30, 2006, 09:32:23
Job time : 498 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 30, 2006, 09:13:58 ; Search time 148 Seconds
(without alignments)
1273.118 Million cell updates/sec

Title: US-10-715-129-2

Perfect score: 533

Sequence: 1 MDVGRNRRKMMKMFRRSPKES.....LAFMPSHSNANDVTGCFM 106

Scoring table: BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1p
-Q=/cgn2_1/USPRO.spool/p/US10715129/runat_27012006_154745_14592/app_query.fasta_1.263
-DB=issued_patents_NA -OPMT=fastcap -SUPPLX=p2n.rml -MINMATCH=0.1 -LDOPTC=0
-LDOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -POCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTENT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US10715129 @CGN_1_1_290 @runat_27012006_154745_14592 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -DEV TIMEOUT=120 -WARN TIMEOUT=30
-THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7 -YGAPOP=10
-YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 95.5 | 17.9 | 312957 | 3 | US-09-949-001-31 |
| 2 | 95.5 | 17.9 | 312972 | 3 | US-09-949-001-34 |
| 3 | 95 | 17.8 | 5340 | 3 | US-09-535-145-1 |
| 4 | 95 | 17.8 | 44166 | 3 | US-09-949-016-15829 |
| 5 | 95 | 17.8 | 54180 | 3 | US-09-949-016-14894 |
| 6 | 93.5 | 17.5 | 152582 | 3 | US-09-949-016-12086 |
| 7 | 93.5 | 17.5 | 152583 | 3 | US-09-949-016-17390 |
| 8 | 93.5 | 17.5 | 152583 | 3 | US-09-949-016-17391 |
| 9 | 92 | 17.3 | 601 | 3 | US-09-949-016-115267 |

| | | | | | | |
|------|------|------|--------|---|----------------------|--------------------|
| c 10 | 92 | 17.3 | 601 | 3 | US-09-949-016-144537 | Sequence 144537, A |
| c 11 | 92 | 17.3 | 165651 | 3 | US-09-949-016-13032 | Sequence 13032, A |
| c 12 | 90 | 16.9 | 601 | 3 | US-09-949-016-180704 | Sequence 180704, A |
| c 13 | 90 | 16.9 | 24979 | 2 | US-08-147-777-3 | Sequence 3, Appl1 |
| c 14 | 90 | 16.9 | 24979 | 3 | US-08-452-872-3 | Sequence 3, Appl1 |
| c 15 | 90 | 16.9 | 24979 | 6 | PCT-US93-03985-3 | Sequence 3, Appl1 |
| c 16 | 90 | 16.9 | 70770 | 3 | US-09-949-016-16938 | Sequence 16938, A |
| c 17 | 89.5 | 16.8 | 3870 | 3 | US-09-949-016-5555 | Sequence 5555, Ap |
| c 18 | 89.5 | 16.8 | 3901 | 3 | US-09-949-016-448 | Sequence 448, App |
| c 19 | 89.5 | 16.8 | 16885 | 3 | US-09-949-016-17297 | Sequence 17297, A |
| c 20 | 89.5 | 16.8 | 16885 | 3 | US-09-949-016-12190 | Sequence 12190, A |
| c 21 | 87.5 | 16.4 | 601 | 3 | US-09-949-016-50372 | Sequence 50372, A |
| c 22 | 86.5 | 16.2 | 234884 | 3 | US-09-949-016-16420 | Sequence 16420, A |
| c 23 | 85.5 | 16.0 | 18798 | 3 | US-09-949-016-14339 | Sequence 14339, A |
| c 24 | 85.5 | 16.0 | 85850 | 3 | US-09-949-016-13424 | Sequence 13424, A |
| c 25 | 85 | 15.9 | 247781 | 3 | US-09-949-016-14193 | Sequence 14193, A |
| c 26 | 85 | 15.9 | 767677 | 3 | US-09-949-016-12147 | Sequence 12147, A |
| c 27 | 85 | 15.9 | 767677 | 3 | US-09-949-016-17361 | Sequence 17361, A |
| c 28 | 84.5 | 15.9 | 2093 | 3 | US-09-620-312D-999 | Sequence 999, App |
| c 29 | 84.5 | 15.9 | 2358 | 3 | US-09-248-786A-6328 | Sequence 6328, App |
| c 30 | 84.5 | 15.9 | 18955 | 3 | US-09-949-016-13343 | Sequence 13343, A |
| c 31 | 84.5 | 15.9 | 30678 | 3 | US-09-949-016-12818 | Sequence 12818, A |
| c 32 | 84 | 15.8 | 15252 | 3 | US-09-949-016-13584 | Sequence 13584, A |
| c 33 | 84 | 15.8 | 387902 | 3 | US-09-949-016-14543 | Sequence 14543, A |
| c 34 | 84 | 15.8 | 421883 | 3 | US-09-949-016-12557 | Sequence 12557, A |
| c 35 | 83.5 | 15.7 | 601 | 3 | US-09-949-016-117038 | Sequence 117038, A |
| c 36 | 83.5 | 15.7 | 1012 | 3 | US-09-775-398-66 | Sequence 66, Appl |
| c 37 | 83.5 | 15.7 | 3466 | 3 | US-10-104-047-536 | Sequence 536, App |
| c 38 | 83.5 | 15.7 | 3494 | 3 | US-09-949-016-4323 | Sequence 4323, App |
| c 39 | 83.5 | 15.7 | 5316 | 3 | US-09-814-915A-26 | Sequence 26, Appl |
| c 40 | 83.5 | 15.7 | 194937 | 3 | US-09-949-016-17032 | Sequence 17032, A |
| c 41 | 83.5 | 15.7 | 194937 | 3 | US-09-949-016-17033 | Sequence 17033, A |
| c 42 | 83.5 | 15.7 | 212139 | 3 | US-09-949-016-16065 | Sequence 16065, A |
| c 43 | 83.5 | 15.7 | 325034 | 3 | US-09-949-016-14957 | Sequence 14957, A |
| c 44 | 83.5 | 15.7 | 389504 | 3 | US-09-949-016-11774 | Sequence 11774, A |
| c 45 | 83 | 15.6 | 2554 | 3 | US-09-799-451-716 | Sequence 716, App |

ALIGNMENTS

RESULT 1
US-09-949-001-31
Sequence 31, Application US/09949001
Patent No. 662536
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 312957
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)..(312957)
OTHER INFORMATION: n = A,T,C or G
US-09-949-001-31

Alignment Scores:
Pred. No.: 69.5
Score: 95.50
Percent Similarity: 57.1%
Best Local Similarity: 42.9%
Query Match: 17.9%
DB: 3
Length: 312957
Matches: 27
Conservative: 9
Mismatch: 25
Indels: 2
Gaps: 1

```
US-10-715-129-2 (1-106) x US-09-949-001-31 (1-312957)
Qy 32 LyaArgMetHelProAlaAntyRPheserLeuGluSerLeuPheLeuValGly 51
Db 197983 CGGGCTGTGCTCTCCCAACTCTCTCATATTCTTCTCTCTCTCTCTCTCTCTCTCC 198042
Qy 52 LeuThrAlaSerLeuValLeuProLeuValLeuProLeuProLeuProLeuProLeu 71
Db 198043 CTCTCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 198099
Qy 72 MetLeuLeuValProLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 91
Db 198100 CCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 198159
Qy 91 oSerSer 93
Db 198160 TTCTTCC 198166

RESULT 2
US-09-949-001-34
; Sequence 34, Application US/0949001
; Patent No. 6825336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1000789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 312972
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(312972)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-001-34

Alignment Scores:
Pred. No.: 69.5 Length: 312972
Score: 95.50 Matches: 27
Percent Similarity: 57.1% Conservative: 9
Best Local Similarity: 42.9% Mismatches: 25
Query Match: 17.9% Indels: 2
DB: 3 Gaps: 1

US-10-715-129-2 (1-106) x US-09-949-001-34 (1-312972)
Qy 32 LyaArgMetHelProAlaAntyRPheserLeuGluSerLeuPheLeuValGly 51
Db 197997 CGGGCTGTGCTCTCCCAACTCTCTCATATTCTTCTCTCTCTCTCTCTCTCTCTCC 198056
Qy 52 LeuThrAlaSerLeuValLeuProLeuValLeuProLeuProLeuProLeuProLeu 71
Db 198057 CTCTCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 198113
Qy 72 MetLeuLeuValProLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 91
Db 198114 CCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 198173
Qy 91 oSerSer 93
Db 198174 TTCTTCC 198180

RESULT 3
US-09-535-145-1/c
; Sequence 1, Application US/09535145
; Patent No. 6703220
```

```
; GENERAL INFORMATION:
; APPLICANT: German, Michael S.
; TITLE OF INVENTION: HUMAN NEUROGENIN 3-ENCODING NUCLEOTIDE
; FILE REFERENCE: UCSF-129
; CURRENT APPLICATION NUMBER: US/09/535,145
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3022)..(3666)
; OTHER INFORMATION: Coding sequence of human neurogenin3
US-09-535-145-1

Alignment Scores:
Pred. No.: 0.201 Length: 5340
Score: 95.00 Matches: 36
Percent Similarity: 51.5% Conservative: 17
Best Local Similarity: 35.0% Mismatches: 25
Query Match: 17.8% Indels: 26
DB: 3 Gaps: 6

US-10-715-129-2 (1-106) x US-09-535-145-1 (1-5340)
Qy 5 ArgAsnAsnArgLys-----AsnMetSerPheArgSerProGluLysSerLysGln 22
Db 418 AGGCAAAACAGAGAGACCCCAAAATGCAGAAATAGAGAGAGTCTGAGCGGCGTCA 359
Qy 23 GluLeuArgArgSerPheSerAlaGlnLysArgMetHelProAlaAntyRPheser 42
Db 358 -----ATTAGAAAGAAATATATGATGCCA-----TTTCT 329
Qy 43 LeuGluSerLeuPheLeuValGlyLeuThrAlaSerLeuLeuLeuProLeuVal 62
Db 328 AGGCAAGGGAGAACCACTGTCATTCAAGCTGCTG-CTGCTGCTGCTTCT- 277
Qy 63 LeuProLeuProLeuProLeuProLeuProLeuProLeuProLeuProLeuProLeu 82
Db 276 ---CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 229
Qy 83 LeuLeuValValLeuAlaPhe-----MetProSerSerHisSer 95
Db 228 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 169
Qy 96 AsnAlaAsn 98
Db 168 TTGCTCTCA 160

RESULT 4
US-09-949-016-15829
; Sequence 15829, Application US/0949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 15829
LENGTH: 44166
TYPE: DNA
ORGANISM: Human
US-09-949-016-15829

Alignment Scores:
Pred. No.: 4.51 Length: 44166
Score: 95.00 Matches: 29
Percent Similarity: 51.8% Conservative: 15
Best Local Similarity: 34.1% Mismatches: 33
Query Match: 17.8% Indels: 8
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x US-09-949-016-15829 (1-44166)

Qy 14 ArgSerSerProGluYSerLySerGlnGluLeuArgSerPheSerAlaGlnLyArg 33
Db 4671 AGGTTTCTCTCAGACACACACCTCCCAAGTACCTGACTTTT-----AAAAATGT 4724
Qy 34 MetMetIleProAlaSerTyPheSerLeuGluSerLeuPheLeu-----Val 50
Db 4725 TTCTTTGTTCTCTCTGTTCTTCTTCTATCTATGTTGTTCTTCTCTCTCTCTCT 4784
Qy 51 GlyLeuThrAlaSerLeuLeuLeuProLeuValLeuProProLeuProProPro 70
Db 4785 GTTGTGTATCTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4844
Qy 71 PheMetLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 90
Db 4845 CT 4895
Qy 91 ProSerSerHisSer 95
Db 4896 TCT 4910

RESULT 5

US-09-949-016-14894
Sequence 14894, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14894
LENGTH: 54180
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(54180)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14894

Alignment Scores:
Pred. No.: 6.09 Length: 54180
Score: 95.00 Matches: 29
Percent Similarity: 51.8% Conservative: 15
Best Local Similarity: 34.1% Mismatches: 33
Query Match: 17.8% Indels: 8
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x US-09-949-016-14894 (1-54180)

Qy 14 ArgSerSerProGluYSerLySerGlnGluLeuArgSerPheSerAlaGlnLyArg 33
Db 14825 AGGTTTCTCTCAGACACACCTCCCAAGTACCTGACTTTT-----AAAAATGT 14878
Qy 34 MetMetIleProAlaSerTyPheSerLeuGluSerLeuPheLeu-----Val 50
Db 14879 TTCTTTGTTCTCTCTGTTCTTCTTCTATCTATGTTGTTCTTCTCTCTCTCTCT 14938
Qy 51 GlyLeuThrAlaSerLeuLeuLeuProLeuValLeuProProLeuProProPro 70
Db 14939 GTTGTGTATCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 14998
Qy 71 PheMetLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 90
Db 14999 CT 15049
Qy 91 ProSerSerHisSer 95
Db 15050 TCT 15064

RESULT 6

US-09-949-016-12086/C
Sequence 12086, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12086
LENGTH: 152582
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(152582)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12086

Alignment Scores:
Pred. No.: 43.5 Length: 152582
Score: 93.50 Matches: 37
Percent Similarity: 52.0% Conservative: 15
Best Local Similarity: 37.0% Mismatches: 29
Query Match: 17.5% Indels: 19
DB: 3 Gaps: 5

US-10-715-129-2 (1-106) x US-09-949-016-12086 (1-152582)

Qy 3 ValGlyArgAlaSerLeuValGlnMetSerPheArgSerProGluYSerLyArg 22
Db 108711 GTTAAATCCACACCTCAAGAGTACACATGAGGCTCTCCCAATTAAGCT---ACT 108655
Qy 23 GluLeuArgSerPheSerAlaGlnLyArgMetIleProAlaSerTyPheSer 42
Db 108654 TCCATGTTTAAAGATAGGCTCTCTCAAGAG-----AACTTCTCTCC 108610
Qy 43 LeuGluSerLeuPheLeuLeuValGlyLeuThrAlaSerLeu-----Le 57
Db 108609 TTCTTCT 108556


```

; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-115267

Alignment Scores:
Pred. No.: 0.0195 Length: 601
Score: 92.00 Matches: 28
Percent Similarity: 51.8% Conservative: 15
Best Local Similarity: 33.7% Mismatches: 32
Query Match: 17.3% Indels: 8
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x US-09-949-016-115267 (1-601)

Qy 14 ArgSerProGluYSerLySgInGluLeuArgSerPheSerAlaGlnLyArg 33
Db 236 AGGTTTCTCTCATGACACACACCTCCCAAGTACGACTTTT-----AAAAATGT 183

Qy 34 MetMetLeuProAlaAntyrrPheSerLeuGluSerLeuPheLeu-----Val 50
Db 182 TTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 123

Qy 51 GlyLeuThrAlaSerLeuLeuLeuLeuProLeuValLeuProProLeuProPro 70
Db 122 GTTGTCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63

Qy 71 PheMetLeuLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 90
Db 62 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 12

Qy 91 ProSerSer 93
Db 11 TCCTCTCTCC 3

RESULT 10
US-09-949-016-144537/c
; Sequence 144537, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 144537
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-144537

Alignment Scores:
Pred. No.: 0.0195 Length: 601
Score: 92.00 Matches: 28
Percent Similarity: 51.8% Conservative: 15
Best Local Similarity: 33.7% Mismatches: 32
Query Match: 17.3% Indels: 8
DB: 3 Gaps: 3

US-10-715-129-2 (1-106) x US-09-949-016-144537 (1-601)

Qy 14 ArgSerProGluYSerLySgInGluLeuArgSerPheSerAlaGlnLyArg 33
Db 236 AGGTTTCTCTCATGACACACACCTCCCAAGTACGACTTTT-----AAAAATGT 183
```

```

Qy 34 MetMetLeuProAlaAntyrrPheSerLeuGluSerLeuPheLeu-----Val 50
Db 182 TTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 123

Qy 51 GlyLeuThrAlaSerLeuLeuLeuProLeuValLeuProProLeuProPro 70
Db 122 GTTGTCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63

Qy 71 PheMetLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 90
Db 62 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 12

Qy 91 ProSerSer 93
Db 11 TCCTCTCTCC 3

RESULT 11
US-09-949-016-13032
; Sequence 13032, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13032
; LENGTH: 165651
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(165651)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13032

Alignment Scores:
Pred. No.: 76.4 Length: 165651
Score: 92.00 Matches: 33
Percent Similarity: 49.5% Conservative: 15
Best Local Similarity: 34.0% Mismatches: 25
Query Match: 17.3% Indels: 24
DB: 3 Gaps: 5

US-10-715-129-2 (1-106) x US-09-949-016-13032 (1-165651)

Qy 15 SerSerProGluYSerLySgInGluLeuArgSerPheSerAlaGlnLyArgMet 34
Db 116094 TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 116141

Qy 35 MetLeuProAlaAntyrrPheSerLeuGluSerLeuPheLeuValGlyLeuThrAla 54
Db 116142 -----TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 116180

Qy 55 SerLeuLeuLeuProLeuValLeuProProLeuProProPhe-----Met 72
Db 116181 -----CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 116231

Qy 73 LeuLeuLeuValProIleGlyIleMetValLeuValLeuAlaPheMet 86
Db 116232 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 116291

Qy 87 -----LeuAlaPheMetProSerSerHisSerAlaAntyrrPheVal 101
```

```

Db      116292 CAACCCCTCCTCCTCCTCCTCCTCCTCAACTCCTCAACTCTTCTTTCTCT 116342

RESULT 12
US-09-949-016-180704
; Sequence 180704, Application US/09949016
; Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 180704
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-180704

Alignment Scores:
Pred. No.:          0.0352           Length:        601
Score:             90.00            Matches:         25
Percent Similarity: 47.6%            Conservative:    15
Best Local Similarity: 29.8%          Mismatches:     18
Query Match:       16.9%              Indels:          26
DB:                3                  Gaps:            2

US-10-715-129-2 (1-106) x US-09-949-016-180704 (1-601)

Qy      34 MetMetIleProAlaasnTyRPheserLeugIsurLeuPheuleValGlyLeuthr 53
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      46 CTAGCGTCGCCTAATCATTAACAATACTAACATTACTACTAAATTAACATAACTACTA 105
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy      54 AlaSerLeuLeuile---LeuProLeuValLeuProPro----- 65
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      106 TTGTCACTACTATTATCACCTACCATTAGAATACACCACTATCATCATCATCACTA 165
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy      66 -----LeuPro 67
      ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|

Db      166 CTAATACATATCACTTACCCTACTATTACCACTTACCAAGACTATCATTTACTACTACCA 225

Qy      68 PropProPheMetLeuLeuValProLeglYlleMetValLeuValLeu 87
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      226 CCTCACCATTTACTACTACTAATATGATATCATTTACTACTACTGGTCTTATTAATTA 285
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy      88 AlapneMetpro 91
      :::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db      286 CTACTCTCTACCA 297

RESULT 13
US-08-147-777-3
; Sequence 3, Application US/08147777
; Patent No. 5914265
GENERAL INFORMATION:
APPLICANT: ROOP, Dennis R.
APPLICANT: Rochnagel, Joseph A.
APPLICANT: Greenalgh, David A.
APPLICANT: Yuspa, Stuart H.
TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 611 West Sixth Street
```

```

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/147,777
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: two
APPLICATION NUMBER: 07/876,289
FILING DATE: April 30, 1992
APPLICATION NUMBER: Unassigned (204/144)
FILING DATE: October 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,337
REFERENCE/DOCKET NUMBER: 204/153
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24979 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-147-777-3

Alignment Scores:
Pred. No.:      8.5          Length:    24979
Score:         90.00       Matches:     26
Percent Similarity: 53.9%   Conservative: 15
Best Local Similarity: 34.2% Mismatches:    9
Query Match:    16.9%     Indels:      26
DB:              Gaps:        3

US-10-715-129-2 (1-106) x US-08-147-777-3 (1-24979)
QY      41 PheserienuGluserleuPhelenuValGlyleuthralaserleunuIleupro 60
      ++++++:::|||||++|||+:
Db      6874 TGCGACAGAAATGATTAATTCCTTCTT-----CTTGCCCTTTCTC 6918

QY      61 leuValenuProProlenuProproProPhemeHerleuenuValProilegly--- 79
      |||+|||||+|||||+|||||+|||||+|||||+|||||+
Db      6919 CTTGCCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6972

QY      80 -----llemet 81
      :
Db      6973 CTTCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7032

QY      82 ValleunValValleualaphemetProserSerHisSerbanala 97
      ::|||+|||||+|||||+|||||+|||||+|||||+
Db      7033 CTTCTCTCTCTCTCTCTCTCTCAATCTTCAATCTTCAATCTTCAATCTTCTCT 7080

RESULT 14
US-08-452-872-3
; Sequence 3, Application US/08452872
; Patent No. 6057298
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.
; APPLICANT: Rothnagel, Joseph A.
; APPLICANT: Greenhalgh, David A.
; APPLICANT: Yusepa, Stuart H.
; TITLE OF INVENTION: KERATIN KI EXPRESSION VECTORS
```


THIS PAGE BLANK (uspto)